DOCUMENT RESUME

ED 056 076 TM 000 880

TITLE	Summary of Research Projects 1970-1971.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	R-No-51
PUB DATE	Sep 71
NOTE	215p.
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Behavior; Cognitive Development; Computer Science; Creativity Research; Cultural Differences; Disadvantaged Youth; Early Childhood Education; *Educational Research; Evaluation; Higher Education; *Human Development; Learning; Personality; Primary Education; *Psychological Studies; Psychometrics; *Research Methodology; Research Projects; *Research Reviews (Publications); Secondary Education; Social Behavior; Testing; Vocational Education

ABSTRACT

Reviews research carried on at Educational Testing Service during the period July 1, 1970 through June 30, 1971. The research areas covered are: (1) Psychological Research (human development, learning and cognition, personality, and social behavior); (2) Educational Applications (early and primary education, secondary education, higher education, graduate and professional education, cultural diversity and the disadvantaged, career and vocation, and evaluation of educational systems); and (3) Theories, Methods and Systems Research (psychometric theory, methodological research and computer studies). (CK)

ED056076

SUMMARY OF RESEARCH PROJECTS

1970 - 1971

(Report #51)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Educational Testing Service
Princeton, New Jersey
September 1971

TM 000 880

ED056076

FOREWORD

The "Summary of Research Projects" reports on research carried on at ETS during the period July 1, 1970, through June 30, 1971.

Questions concerning the studies reported herein should be directed to the authors. If a report on a study is available, this will be indicated by a number appearing in the body of the write-up. When requesting a copy of a report please use this number.

TABLE OF CONTENTS

I. PSYCHOLOGICAL RESEARCH

A. Human Development

Baker, A. H.

Developmental Changes in Self/World Differentiation 173

Developmental Model of Cognitive Style 122

Ban, P.

Mothers and Fathers, Girls and Boys: Attachment Behavior in the One-Year-Old (with M. Lewis) 178

Baskett, G. D.

Learning of an Alternating Stimulus Pattern in Young Children (with M. Lewis) 156

Baumel, M.

Infants' Attentional Distribution across Two Modalities (with M. Lewis) 115

Darley, S.

Activity and the Cardiac Response (with M. Lewis) 164

Emmerich, W.

Differentiation and Development of Social Norms 135

Freedle, R.

On Relating an Infant's Observation Time of Visual Stimuli with Choice Theory Analysis (with M. Lewis) 161

Lewis, M.

Attentional Distribution as a Function of Novelty and Incongruity (with C. Wilson, N. Van Wyk, and J. Weisman) 155

Developmental Study of Attention within the First Two Years of Life (with C. Wilson, M. Baumel, P. Ban, and N. Van Wyk) 114

Lewis, M. (Continued)

A Developmental Study of the Cardiac Response to an Auditory Signal: Onset and Offset of a Stimulus (with P. Ban and M. Baumel) 165

Infant Development in Lower-Class American Families (with P. Sarett) 176

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development (with C. Wilson) 157

Mother-Infant Interaction and Infant Development among the Wolof of Senegal (with D. Lusk) 177

The Mother-Infant Interaction in Africa (with D. Lusk) 160

Social Class and Sex Differences in the Attachment and Play Behavior of the Year-Old Infant 162

Stability of Attachment Behavior: A Transformational Analysis (with P. Ban) 180

State as an Infant-Environmental Interaction: Mother-Infant Behavior (with P. Sarett, Y. Watson, and C. Wilson) 117

What's Thrown Out with the Bath Water: A Baby? (with N. Johnson) 163

Wilson, C.

A Developmental Study of Attention: A Multivariate Approach (with M. Lewis) 112

Wohlwill, J.

A Longitudinal Study of the Formation of Concepts of Conservation and Classification, as Related to the Development of Spontaneous Cognitive Activities 120

Methodology and Research Strategy in the Study of Developmental Change 121

B. Learning and Cognition

Cahen, L.

Difficulty Factors in Learning to Spell 169

Generalizability of Cognitive Styles 152

Carroll, J.

Application of the Lognormal Model of Word Frequency Distribution 149

Carroll, J. (Continued)

Factors in Comprehension of Verbal Material (with R. Freedle) 147
International Study of Educational Attainment in French as a Foreign Language 185
U. S. National Testing in French as a Foreign Language 185
Learning from Verbal Discourse in Educational Media 111
School-Age Children's Comprehension of Words with Multiple Grammatical Functions 118
Simulation of Learning Processes in a Restricted Set of Experiments from the Literature 148
Studies of the Acquisition of Grammar in a Foreign Language 151

Freedle, R.

Extensions of Luce's Choice Theory 129
Human Choice Behavior 153
Human Choice Behavior in Psycholinguistic Studies 167
Informational Redundancy in Messages as a Function of Stimulus Definition and Complexity (with P. Kingsley) 103

Fremer, J.

Recall of Approximations to English 186

Hale, G.

Developmental Studies of Children's Incidental Learning and Distractibility 166
Interrelations in Children's Learning of Verbal and Pictorial Paired Associates 170
Selective Attention in Children 119

Harman, H.

Reference Measures for Cognitive and Personality Factors 67

Jacobs, P.

Training of Cognitive Functions in Children 144

Ward, W.

Labeling and Memory in Children 100

C. Personality

Alker, H.

Are Bayesian Optimal Decisions Artificially Intelligent? (with M. Hermann) 146

Damarin, F.

Leveling-Sharpening as a Cognitive Style (with V. McKenna and S. Messick) 78

The Psychology of Response Styles (with S. Messick) 130

Study of Objective Tests of Personality in Adults 97

Frederiksen, N.

Learning to Formulate Original Responses in Relation to Anxiety and Defensiveness (with F. Evans) 145

Freedle, R.

Stimulus Clues in the Judgment of Veracity (with D. Baskett and P. Kingsley) 168

Kogan, N.

Longitudinal Study of Creativity 134

Messick, S.

Cognition of Affect 137

Cognitive Styles and Attitudinal Consistency 142

Cognitive Styles and Psychopathology 133

Dimensions of Perceived Item Similarity 77

Judgmental Dimensions of Psychopathology 77

Perceptual Factors in Cognition and Personality (with J. French) 68

Personality Organization in Cognitive and Social Processes 79

Stricker, L.

Response Styles and Second-Order Factors in the Personality Domain 98

Retest Changes on Personality Scales 88

Ward, W.

Creativity and Classroom Behavior 131

Ward, W. (Continued)

Convergent and Divergent Measurement of Creativity in Children 131

A Field Study of Nonverbal Creativity 102

Growth and Stability of Creativity During the School Year 132

Motivation and Capacity in Children's Creativity (with N. Kogan) 136

D. Social Behavior

Baskett, G. D.

Perceptual Errors as a Function of Sex-Role Attitudes (with S. Darley, M. Keiffer, and P. Kingsley) 104

Carlson, J.

Cross-Cultural Study of Group Risk Taking 99

Keiffer, M.

Bargaining in the Mexican Marketplace 94

Discrimination Experienced by Academic Female Psychologists 94

The Effect of a Bargainer's Role on Bargaining Behavior 95

Women's Status and Population Control: A Bibliography (with P. Warren) 96

Kogan, N.

Determinants of Risk Taking for Others (with A. Teger) 140

Dimensions and Categories of Physiognomic Judgment 139

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (with M. Hermann) 89

Group Decision-Making in a Minimal Social Situation 101

Individual and Group Decision-Making under Conditions of Responsibility for Others 90

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts 141

Kogan, N. (Continued)

Negotiation Behavior under Conditions of Observation and Consultation 91

Pessimism-Optimism and Risk-Taking 92

Risk Taking in the Context of Intergroup Negotiation 92

Stricker, L.

Dimensions of Conformity (with S. Messick) 138

II. EDUCATIONAL APPLICATIONS

A. Early and Primary Education

Ball, S.

Preschool Television for Children 40

Cahen, L.

Manipulated Expectancies and Teacher Bias 171

Chittenden, E.

Assessment and Development of Cognitive Skills in Young Children: A Classroom Approach (with A. Bussis, J. Orost, and M. Tanaka) 192

Development of Procedures for Assessing Open Education in Follow Through Programs (with M. A. Amarel, A. Bussis, and M. Tanaka) 193

Study of Teacher Change in Open Educational Settings (with A. Bussis, and M. A. Amarel) 194

Ironside, R.

IMS Advancement Tests 23

Marco, G.

Michigan Title I Evaluation (with J. Weisbrodt) 20

Massad, C.

A Study of the Interrelationships among Reading Comprehension Ability, Writing Ability, and the Linguistic Structure of the Written Language of Sixth-Grade Children 187

Norton, D.

Minneapolis Pyramid Sibling Study 19

B. Secondary Education

Angoff, W.

PAA-SAT Equating Study 14

Study to Relate the Scales for the Canadian SAT (English) and CEEB SAT 16

Diederich, P.

Newspaper Tests 175

Tests of Cultural Development 93

Evans, F.

Mathematics Coaching Study (with L. Pike) 15

Hilton, T.

Analysis of the Stability and Instability of Student Growth (with A. Beaton and C. Bower) 199

Growth as a Function of the Amount of Training (with G. Berglund) 63

A Study of Academic Prediction and Growth in Grades 5 through 12 (with C. Werts) 59

Katz, M.

PSAT-AIM Norming and Follow-up Study 11

Modu, C.

American History and Social Studies Essay Study 8

Pitcher, B.

College Board Activity and Participation Statistics 6

Thompson, R.

Curricular Appropriateness of the College Board Science Achievement Tests 7

C. Higher Education

Angoff, W.

TOEFL Language Study (with A. Sharon) 35

Ball, S.

Motivation, Teaching, and Learning in the Classroom 47

Centra, J.

College Environments and Student Academic Achievement 36

Evans, F.

Development of Undergraduate Rating Scales (with T. Hilton) 31

Peterson, R.

Cambodia Aftermath Survey 189

Development of an Institutional Goals Inventory, Phase 1 22

Schrader, W.

Quantitative Comparison Validity Study 9

Sharon, A.

Academic Credit by Examination 190

Remediation in Junior College 18

Warren, J.

Assessment of Junior College Student Concerns (Student Reactions to College) 13

Study of Academic Performance 188

D. Graduate and Professional Education

Boldt, R.

ATGSB Grade Grouping Study 52

Burns, R.

Investigation of Graduate Admissions Policies and Procedures 32

Carlson, A.

ATGSB New Item Type Research 53

LSAT Factor Analysis Validity Study 27

Crooks, L.

Criterion Study 51

In-Basket Study 54

Echternacht, G.

Alternate Methods of GRE Advanced Tests 33

Ironside, R.

University of Puerto Rico Interim Project for Physician Performance Criteria 25

University of Puerto Rico Literature Study 21

Pitcher, B.

A Comparison of the Prediction of First-Year Average Grades in Graduate Schools of Business for Interrupted and Uninterrupted Students 57

Indicators of College Quality as Predictors of Success in Graduate Schools of Business (with W. Schrader) 58

Schrader, W.

Discriminant Function Analysis of Law School Performance (with B. Pitcher) 29

Summary of Validity Studies of Experimental Item Types, 1947-1967 30

Sharon, A.

The Moderating Effect of TOEFL on ATGSB 55

The Moderating Effect of TOEFL on GRE 34

E. Cultural Diversity and the Disadvantaged

Angoff, W.

PSAT Item-Bias Study 10

Ball, S.

The Continuing Evaluation of Sesame Street 43

Evaluation of Selected Aspects of the CTW Utilization Effort 43

Campbell, J.

Test and Job Performance of Various Subgroups within Specified Occupations (with L. Crooks, D. Rock, M. Mahoney, and M. E. Parry) 198

Emmerich, W.

Personality Structure and Development in Headstart Children 38

Evans, F.

ATGSB Speededness Study (with R. Reilly) 56

LSAT Speededness Study (with R. Reilly) 28

Hilton, T.

Home Environment and Educational Development (with D. Lantz and C. Patrick) 61

Hsia, J.

Evanston Integration 184

Shipman, V.

Disadvantaged Children and Their First School Experiences: A Longitudinal Study 37

Stricker, L.

Social Stratification 143

Warren, J.

Chicano Validity Study 17

F. Career and Vocation

Crooks, L.

The Executive Study 65

In-Basket Development and Research 66

Goldstein, L.

Perspectives on Law Enforcement:
1. Characteristics of Police Applicants 41

Supplementary Analysis of Police Applicant Data 48

Harrell, T.

Selection of MBA Candidates for Later Job Effectiveness 50

Hilton, T.

Education After High School in 1961 and 1967 (with C. Bower) 60

A Study of Intellectual Growth and Vocational Development (with P. Casserly, F. Evans, N. Freeberg, W. Godwin, M. Katz, M. Patton, D. Rock, M. Rosenfeld, and J. Warren) 195

Katz, M.

System of Interactive Guidance and Information 191

G. Evaluation of Educational Systems

Anderson, S.

National Assessment (with A. Barron) 39

Cahen, L.

Matrix Sampling and Abilities 172

Ironside, R.

EIP Assessment of Impact 24

Marco, G.

Chicago Title I Evaluation 45

Educational System Information Program--Cincinnati 45

The Relationship of School District Wealth to Student Achievement and to Other School Characteristics 1

Seibel, D.

Educational Program Audits 44

Trismen, D.

Taxonomy and Survey Project 46

III. THEORIES, METHODS, AND SYSTEMS RESEARCH

A. Psychometric Theory

Bechtel, G.

Least Squares Matrix Factorization and Multidimensional Scaling 110

13

Jöreskog, K.

Exploratory Factor Analysis (with M. van Thillo) 107

A General Method for Analysis of Covariance Structures 86

A General Method for Estimating a Linear Structural Equation System (with M. van Thillo) 87

Kristof, W.

Factorization of Covariance Matrices in Terms of a Specified Pattern or Structure (with B. Wingersky) 182

Generalization of the Orthogonal Procrustes Rotation Procedure to More Than Two Matrices (with B. Wingersky). 182

On Proportional Profiles in Factor Analysis 126

On the Sampling Theory of Reliability Estimation 76

Some General Results on Fit in Factor Rotation 183

On a Statistic Arising in Testing Correlation 127

A Theorem on the Trace of Certain Matrix Products and Some Applications 128

On the Theory of a Set of Tests Which Differ Only in Length 128

Lord, F.

Accuracy of Estimated True Scores (with M. Hamilton) 69

Disattenuated Partial Correlation 70

Efficiency of Estimation When There Is Only One Common Factor (with M. Wingersky) 71

Estimating Item Characteristic Curves (with M. Wingersky) 71

Estimating Item Parameters (with M. Wingersky) 71

Strong True-Score Theory (with M. Wingersky and M. Hamilton) 73

Tailored Testing 74

Two-Stage Tests (with M. Hamilton) 75

Novick, M.

The Choice of Prior Probabilities for Bayesian (Logical Probability) Statistical Analysis 125

Rock, D.

Advanced Moderator Variable Study (with F. Evans) 26

B. Methodological Research

Benjamin, R.

Longitudinal Study of Item Responses 62

Donlon, T.

Differential Weighting of Item Distracters in Data Sufficiency Items (with G. Dunteman) 12

Gulliksen, H.

Intercultural Attitude Study 105

Learning 106

Harman, H.

Planning, Exploration and Adaptation of Statistical and Psychometric Techniques 5

Lord, F.

The Self-Scoring Flexilevel Test 72

Marco, G.

Cognitive and Affective Dimensions of Differential Achievement on Open and Closed Book Achievement Tests 49

Trismen, D.

Models of Change and Response Uncertainty 46

Wingersky, B.

Data Analysis Assuming Only Random Sampling 3

C. Computer Studies

Beaton, A.

Multivariate Change Pilot Study (with T. Hilton) 64

Harman, H.
    Factor Analysis 2

Lavine, L.
    Model for ETS Shipping 4
    Scheduling Techniques 4

Leaf, W.
    A Computer Model of Everyday Human Thinking 108
    A List Processing Extension to PL-I 109

The Relationship of School District Wealth to Student Achievement and to Other School Characteristics (1-16) Mr. Marco

PURPOSE: To relate cost per pupil, property valuation, and other indicators of N. J. school district wealth to school process and output.

PROGRESS: SAT data and school data from the *Annual Report of the Commissioner of Education* and other sources are being collected. Correlations and scatterplots will be produced and interpreted.

Factor Analysis (29-40) Mr. Harman

PURPOSE: Research and development work and computer programming related to factor analysis.

PROGRESS: Review of factor analysis literature is a continuing activity. An attempt is made at complete coverage of theoretical developments. Applications of factor analysis are well documented in the psychological literature so there is no necessity for separate review. Applications in other fields, as varied as sociology and meteorology, political science and medicine, geography and business, is well sampled--the total number of references exceeding 200 titles.

The project director gave an invited address at the International Symposium on Yeasts as Models in Science and Techniques, Smolenice Castle, Czechoslovakia entitled "How Factor Analysis Can Be Used in Classification." It makes use of the minres method of factor analysis and groups strains of yeast according to their factor profiles. The similarities are judged in the higher-dimensional space determined by the factor analysis using a grouping procedure and an associated validation procedure developed by Bary Wingersky.

Data Analysis Assuming Only Random Sampling (29-41) Mr. Wingersky

PURPOSE: In many specific real world problems we are interested in distinguishing individuals of one type from individuals of another type on the basis of information provided on their attributes, or test scores, and the like, without knowing what types of individuals we have to begin with.

PROGRESS: A criterion for recognizing when a body of data exhibits a high degree of structure has been found. This criterion does not require the specification of the type of structure it would indicate. A procedure for reducing the dimensionality of the space in which the data are located in such a way as to maximize the degree of structure of the sample resulting in the subspace is being developed. For any space or subspace, a program to show whether the structure indicated by the criterion is due to one group with strong multivariate relations, or several groups with possibly different orientations in the space, is completed. The program package consists of two major parts: (1) group procedure and (2) validation procedures including discriminant analysis and classification. These procedures have been applied to such varied areas as classification of yeasts (Harman), classification of graduate schools of business (Pitcher), and grouping of college students on the basis of personality inventory scores (Warren).

Model for ETS Shipping (29-42) Mr. Lavine

PURPOSE: To design and program a simulation model that will permit program directors, analysts, and others to evaluate alternative strategies for shipping test materials.

PROGRESS: This project has been quiescent for most of the year. The simulation program is available, but other pressing requirements made it difficult to carry out further work. The next effort should be a pilot study in one of the ongoing test programs.

Scheduling Techniques (29-43) Mr. Lavine

PURPOSE: Research, development and adaptation of existing programs concerned with general problems of planning and scheduling of time and cost.

PROGRESS: A preliminary PERT Network was prepared to assist in planning several major changes for the CEEB ATP. This effort was not carried further. Assistance in the preparation of PERT-like schedules was provided to two or three other projects in the course of the year.

Continuing problems with IBM's PMS system have plagued us throughout the year. Active support from IBM has been lacking; in recent weeks this situation seems to have changed. We are looking forward to a completely computerized PERT service that will handle both scheduling and costs.

Planning, Exploration and Adaptation of Statistical and Psychometric Techniques (29-61) Mr. Harman

PURPOSE: An important role of the Division of Analytical Studies and Services is to improve the analytical techniques used in support of the operating programs of ETS. This project is designed to provide a continuing effort to explore a variety of statistical and psychometric methods that appear to have potential for more valid or more efficient operations.

One study, being carried out by Mr. Case, involves the investigation of the theoretical nature and methodological possibilities of the Rasch model or one-parameter logistic distribution model as it relates to the following problems:

1. Identification of items not similar to others in a given test
2. Comparison of item difficulties
3. Equating tests with small number of items in common.

Review of current literature and discussion with others working infield (e.g., E. Andersen, F. Lord, B. Wright, etc.) has been pursued. A method of estimating relative item difficulty for pairs of items has been devised; a preliminary computer program was written for comparing these estimates to establish the effectiveness of the model. An average of about five percent discrepancy was found among estimates on some SAT data. A revision of this program is to provide graphical portrayal of the components of the estimates of relative item difficulties in order to identify dissimilar items. Data from Real Estate licensing programs will be used in addition to the SAT data.

Another study is aimed at providing, economically, the factorial structure of a test in the hope that it will lead to improvements both in the planning of future tests and in the actual writing of test items. As a first step, it was necessary to get the best possible index of relationship between any two test items. The tetrachoric correlation appears to be most suitable, but an accurate computing routine was not readily available.

Various methods of calculating tetrachoric r in the past have relied upon the use of series approximations or a combination of series approximations and iterations. Although adequate accuracy was obtained for most values of r, difficulty was still encountered near the limits. Mr. Kirk sought to find an algorithm which would be relatively fast and provide good results for extreme values. A combination of Gaussian quadrature to evaluate the probability integral, supplemented by a Newton-Raphson iterative technique was tried and found to be quite fast and provided good results over the entire range.

MAJOR FINDINGS: The use of a combination of either a 5-point or an 8-point quadrature and an iterative technique will give a rapidly converging estimate of tetrachoric r (RB-71-35).

College Board Activity and Participation Statistics

(CEEB, 130-2)

Miss Pitcher

PURPOSE: This study is concerned with the preparation of selected descriptive statistics on College Board candidates to describe patterns of test participation with particular reference to shifts that may affect score interpretation and program planning.

PROGRESS: Candidates and Tests, 1969-70 was completed and distributed.

MAJOR FINDINGS: Candidate and test volume in the admissions testing program appeared to be leveling off in 1969-70. Four indices of overall testing activity used in the admissions testing program showed the following changes in 1969-70 over 1968-69 volume: (1) an increase of 1.1% in number of Scholastic Aptitude Tests administered, (2) a decrease of 3.0% in number of Achievement Test series, (3) a decrease of 0.04% in total test volume, and (4) an increase of 1.2% in number of registrations. The number of Preliminary Scholastic Aptitude Tests decreased 9.1%.

Curricular Appropriateness of the College Board Science Achievement Tests (CEEB, 130-8) Mr. Thompson

PURPOSE: To determine if the tests are equally appropriate for students of the newer and students of the more traditional high school science courses.

PROGRESS: Six forms of the Biology Test, 12 forms of the Chemistry Test, and eight forms of the Physics Test have been analyzed. Scores on the science tests of students from different courses have been adjusted for the performance of these students on SAT and questions in the science tests rated appropriate for students of the different courses. The two latest forms of each science test were analyzed for samples based on student replies regarding courses studied and for samples based on teacher verifications of student replies.

MAJOR FINDINGS: Results to date indicate that the Biology, Chemistry, and Physics Achievement Tests are equally appropriate for students of the newer and the more traditional high school science courses. These results hold true for students whose replies about the course they studied were verified by their teachers.

Results for the Physics Test are reported by Guenter Schwarz and Raymond E. Thompson in "The College Board Physics Achievement Test," The Physics Teacher, Vol. 7, No. 3, March 1969. A report summarizing all results is in progress.

American History and Social Studies Essay Study Mr. Modu
(CEEB, 130-9)

PURPOSE: The purpose of the study is to determine whether an essay achievement test in American History and Social Studies adds a unique component to the pool of information presently obtained from an all-objective examination of the College Board.

PROGRESS: A 20-minute essay question was administered to over 11,000 candidates as part of the one-hour achievement test in American History and Social Studies at the March 1971 administration of the college entrance examinations. Each essay test was scored by two readers. The answer sheets of a sample of candidates who took both the SAT tests and the achievement tests in American History and Social Studies and English Composition have been rescored to produce several subscores to be used in various analyses which are presently under way.

Quantitative Comparison Validity Study (CEEB, 130-22)

Mr. Schrader

PURPOSE: To obtain evidence on the relative effectiveness of a 30-minute Quantitative Comparison Test and of SAT-Mathematical for predicting college grades.

PROGRESS: Special test administrations were conducted at 13 institutions in the fall of 1970. Rosters prepared on College Board Validity Study Service forms have been forwarded to College Board Regional Offices for use in collecting validity study data.

PSAT Item-Bias Study (CEEB, 130-23) Mr. Angoff

PURPOSE: The purpose of this study is to examine the item x group interaction of the Verbal and Mathematical items of the PSAT given in November of 1970 for various pairs of samples: Unselected Blacks in Atlanta versus (1) a second unselected sample of Blacks in Atlanta; (2) unselected Whites in Atlanta; (3) unselected Blacks in Savannah; and (4) unselected Blacks in towns in Georgia numbering fewer than 25,000. Also (5) Blacks in Atlanta versus matched Whites in Atlanta (matched on Math for studying Verbal; matched on Verbal for studying Math); and (6) unselected Whites in Atlanta versus a second unselected sample of Whites in Atlanta. The comparison indicated in (2) is of the conventional type. The comparison indicated in (3) is intended to determine whether there is an interaction between the items and the particular urban area for Blacks. That in (4) is an item x rural-urban interaction. That in (5) is an item x race interaction when there is matching on an independent ability dimension. The comparisons in (1) and (6) are intended to provide baselines for evaluating the other comparisons.

PROGRESS: The request for data was postponed until January-February, well after all the data for the regular PSAT program had been processed. Item analysis data and plots of deltas for the indicated samples have been received and are presently being reviewed.

PSAT-AIM Norming and Follow-up Study (CEEB, 140-80) Mr. Katz

PURPOSE: The major research objectives of this study include: (1) the development of grade 10 norms for the PSAT, (2) the revision of grade 11 and 12 PSAT norms, and (3) the development of norms and other materials for the use and interpretation of the Academic Interest Measures (AIM).

PROGRESS: Three sequential phases of research were planned. Phase 1 was directed toward the development of norms for both PSAT and AIM and toward gathering evidence of AIM's construct validity; phase 2 was concerned with determining the internal consistencies and one- and two-year stabilities of AIM scale scores; phase 3 centered on investigating the predictive validities of AIM scales.

The final report has appeared in two volumes, as College Board RDR-70-71, No. 4, ETS RB-70-57, and RDR-70-71, No. 5, ETS RB-70-67. Under the overall title, The Measurement of Academic Interests, Part I is "Characteristics of the Academic Interest Measures," and Part II is "The Predictive Validities of Academic Interest Measures." AIM norms are bound separately as Appendix D to Part II.

Norms for PSAT at grade levels 10, 11, and 12 have been published in College Board Score Reports, 1969.

Four journal articles have been prepared on the following topics:

(1) The Contribution of Academic Interest Measures to the Prediction of Marks
(2) Interests and Values: A Study in Structure
(3) The Prediction of Interests
(4) The Discriminating Power of Measures of Academic Interests and Occupational Values.

Differential Weighting of Item Distracters in Data Sufficiency Items (CEEB, 141-5)

Mr. Donlon
Mr. Dunteman

PURPOSE: Differential weighting of wrong-answer responses to multiple-choice items may provide additional information not currently reflected in the conventional scoring methods. Data sufficiency items would seem to offer particular promise in that the meaning of responses is constant for all items regardless of content. A comparison will be made of the efficiency of an a priori weighting system and a system using the judgments of mathematics specialists.

PROGRESS: The experimental tests were prepared and administered to approximately 1800 freshman students at seven colleges. Judged estimates of option merits have been secured, and grade-point averages have been received from the schools. Preliminary analyses of the data revealed that none of a number of empirically or rationally derived composite scores, based upon different linear combinations of the different response category scores, yielded an alternate forms reliability as high as the Total Rights score alone.

A number of additional analyses are being performed. They include determining the composite weights by canonical correlational analysis, principal components analysis, and multiple regression analysis, using the Mathematics Achievement Test as the criterion. Validities are being determined for those linear combinations that have a reliability greater than the Total Rights score reliability.

Assessment of Junior College Student Concerns (Student Reactions to College) (CEEB, 141-10) Mr. Warren

PURPOSE: An instrument is being developed to assess the views of junior college students about their college experiences and the congruence or disjunction they see between their experiences and their reasons for being in college. The instrument is to provide a vehicle for communication between students and staff.

PROGRESS: Interviews with junior college students, faculty members, and administrators, followed by summary conferences, have provided the content and guides for the format of the instrument. About 6,000 students in 24 junior colleges across the country have responded to an experimental version of the questionnaire. Analyses now being conducted are concerned with item characteristics and the identification of student subgroups having different views of college.

PAA-SAT Equating Study (CEEB, 142-5) Mr. Angoff

PURPOSE: The purpose of this study is: (1) to relate the 200-800 scales for the Puerto Rican PAA to the 200-800 scales for the SAT, and (2) to determine, if possible, the nature of the items that fail to have the same psychological meaning in the two language-cultures. The method of the study is to choose sets of items originally appearing in Spanish and others originally in English, to translate each into the other language, and to administer both sets in the appropriate language mode to both types of students. Items showing little item x group interaction would then be chosen and assembled as "common" items comprising equating sections in English and in Spanish. These equating sections would be used to equate the scales in the two languages.

PROGRESS: The items for the study have been translated, English to Spanish and Spanish to English, and reviewed by the College Board Puerto Rico Office. The 10 Spanish-language pretest forms (5 Verbal and 5 Mathematical), requiring 25 minutes of administration time, and 8 English pretest forms (4 Verbal and 4 Mathematical), requiring 30 minutes of administration time, were administered in November 1970. Delta plots of Verbal items and of Mathematical items, as observed for SAT candidates and PAA candidates, have been drawn up. Forty verbal items and 25 mathematical items have been chosen for equating PAA-Verbal to SAT-Verbal and PAA-Math to SAT-Math. The equating sections are presently being assembled and will be administered along with the operational forms of PAA and SAT in November 1971. The final analysis of the equating and the report are expected in the Spring of 1972.

Mathematics Coaching Study (CEEB, 142-9)

Mr. Evans
Mr. Pike

PURPOSE: The study was designed to investigate whether student scores on Quantitative Comparison, Data Sufficiency, and traditional SAT-mathematics items could be substantially increased by coaching combined with short-term instruction.

PROGRESS: Approximately 50 high school junior volunteers from each of 12 schools in New Jersey, Pennsylvania, and Ohio were randomly assigned to one of three treatment groups or a control group. All of these students were tested in mid-October 1970. They took a regular SAT plus a 55-item Quantitative Comparison mathematics test. The students in the treatment groups then received seven weeks of instruction in either quantitative comparisons data sufficiency, or traditional SAT-mathematics items while those in the control group received no special instruction. Special materials for both in-class and out-of-class assignments were developed for each of the three treatment groups. Students in the treatment group were instructed in three-hour sessions on seven consecutive Saturday mornings. At the end of the instructional period (early December 1970) all students (treatment and control) took parallel forms of the tests they had taken in October. Students who were in the control groups were then given an opportunity to receive special instruction similar to that provided treatment group students. All students were given the opportunity to take a regular SAT free of charge in April 1971 (about 90% of the 509 students in the study took advantage of this free testing).

MAJOR FINDINGS: Only preliminary analyses have been completed to date. In general these analyses show that for the students in this study all three mathematics item types are somewhat susceptible to coaching combined with short-term instruction. Instructed groups gained more from pre- to posttest than control groups on the set of items similar to those on which they received instruction. For example, students who received instruction in Quantitative Comparisons items gained more on the 55-item Quantitative Comparisons Test (from pretest to posttest) than students who were instructed in Data Sufficiency or traditional SAT-math or students who received no special instruction. Caution must be exercised in interpreting these results. The students in this study were highly motivated high school juniors who gave up nine consecutive Saturday mornings to receive the special instruction. Unlike so-called SAT cram courses which typically provide 10 or 12 hours of instruction in verbal and mathematics combined, the 21 hours of instruction provided to students in the study was concentrated on one small aspect of the SAT, namely a single SAT mathematics item-type.

Mr. Angoff

Study to Relate the Scales for the Canadian SAT (English) and CEEB SAT (CEEB, 142-11)

PURPOSE: The intent of the study is to determine the equations converting the scales for the College Board SAT-Verbal and SAT-Mathematical to the corresponding scales for the Canadian SAT. To accomplish the conversions, a 30-minute operational section of the January 1971 Canadian SAT-Verbal test will be administered with the January 1971 form of the College Board SAT as a variable ("experimental") section. Similarly for the mathematical test. Correspondingly, a 30-minute operational section of the January 1971 College Board SAT-Verbal will be administered with the January 1971 form of the Canadian SAT as a variable section. Similarly for the mathematical test. The two verbal sections, thus administered to students in both countries, will be used as a combined equating section to relate the two verbal scales. Similarly for the mathematical test.

PROGRESS: Each of the two testing organizations has transmitted to the other the required operational sections, and both organizations have administered these sections as variable sections in the operational SAT and CSAT in January 1971. Data have been received from SACU in the form of tape records of candidates who have taken the operational College Board section as their variable section. Delta plots for both the verbal and mathematical items have been examined, and items have been chosen (75 Verbal and 35 Math) for equating CSAT-Verbal to SAT-Verbal and CSAT-Math to SAT-Math. The equating results should be available by June 30, 1971, after which time a report of the work will be written.

Chicano Validity Study (CEEB, 142-13) Mr. Warren

PURPOSE: Mexican-Americans frequently contend that their bilingual and bicultural background prevents standardized tests from being as valid for them as for members of the majority culture. This contention will be examined by comparing regression equations predicting college performance from test scores and high school grades for Chicano and non-Chicano students in three California state colleges.

PROGRESS: Test scores and high school and college grades are being recorded fror the records of the three participating colleges.

Remediation in Junior College (CEEB, 175-1) Mr. Sharon

PURPOSE: The major objective of this experimental study was to determine how effective are remedial courses in junior colleges--do they in fact prepare the student for subsequent work? A second purpose was to determine how effective are present placement policies and instruments--do they place students in the highest level course they can successfully complete?

PROGRESS: Two Connecticut community colleges have placed some of their students identified for remediation into regular English and math courses on an experimental basis. Comparisons were made on a variety of predictor and criterion data gathered on students who did and did not receive remediation. These included measures of ability, interest, motivation, persistence, and performance. The study has been completed and reported in RDR-70-71 No. 2, RB-70-50.

MAJOR FINDINGS: Some of the remedial English courses considered in this study had a slight but significant effect in raising the grades in the subsequent regular English course. The single math remedial course which was evaluated appears to be more effective than the English remedial courses. The course improved the performance of the students in the subsequent regular course by more than a half of a letter grade. In addition, the students were much more satisfied with it than with the regular course. A fair evaluation of the placement procedures is difficult to make. Since the failure rate in the regular courses is very low, identifying those few who would fail is a formidable task for any placement system. The placement procedures are apparently more effective in assigning students to appropriate math courses than to appropriate English courses.

Minneapolis Pyramid Sibling Study (268-2) Mr. Norton

PURPOSE: The Pyramid Reading Project is a major Title I ESEA effort in the Minneapolis Public Schools. Its impact should appear in reading scores of pupils but is hypothesized to be cumulative and not fully assessible for four or more years. The methodology of a sibling study promises better validity for project reports than other evaluative options. A longitudinal study of program outcomes is to be initiated.

PROGRESS: The project was initiated in June 1971.

Michigan Title I Evaluation (333-30)

Mr. Marco
Mr. Weisbrodt

PURPOSE: To estimate the impact of the Michigan Title I program on third-grade reading achievement.

PROGRESS: Tests were administered in the fall and the spring. Spring tests have yet to be scored. A regression analysis is planned that will identify which programs were the most effective for low-, medium-, and high-scoring students.

University of Puerto Rico Literature Study (345-1)

Mr. Ironside

PURPOSE: To discover what the diverse medical literature reveals concerning (a) specific performance criteria in the practice of general medicine, and (b) procedures which have been used in developing such criteria. A third but lesser purpose was to note implications in the literature for measurement of criterion attainment. All of this is background for continued work with the University of Puerto Rico's Medical School. UPR is offering a refresher course to graduates of foreign medical schools, and ETS has been involved in course evaluation, instrument development, student selection, and matching course content to student needs.

PROGRESS: The search has extended to 150 carefully selected sources (from among hundreds) in six major journals, other journals, books, and reports of research in progress. In December 1970 the first paper was distributed to UPR staff for review and use. A second "edition" is in progress, with additional sources employed, and is expected to appear as a Research Memorandum.

MAJOR FINDINGS: Performance criteria related to quality medical practice have been defined for many years, but only since the middle 1950's have there been serious studies of actual physician performance and application of findings to medical education efforts. Professional organizations (AMA, specialty societies) are now beginning to assess individual physician practice as opposed to the earlier audits which were concerned with the whole hospital. Criteria vary in specificity and range, and the literature reveals only a handful of efforts to define and rate given "behaviors." But a number of important dimensions appear to be of fairly common concern: clinical patient management, maintaining medical knowledge, relating to patients, patient satisfaction, record-keeping, physician responsibilities, and preventive medicine. Study of records has most often been the means of examining performance, but some efforts have been made through observation, discussion of patient-management, simulation tests, and patient interviews.

Development of an Institutional Goals Inventory, Phase 1 (362-1)

Mr. Peterson

PURPOSE: The objective of Phase 1 is to develop a machine scorable Institutional Goals Inventory for use by colleges and universities beginning in the fall of 1971. (Phase 2 is to consist of a national norming project.)

PROGRESS: A series of analyses of data gathered with a preliminary goals inventory in the Carolinas and Virginia were carried out. A revised experimental IGI was prepared consisting of selected goal statements from the preliminary form as well as new items. The new 110-item, 22 scale instrument was administered to samples of 50 to 150 students and faculty at 10 diverse colleges on the West coast during May. Scoring and other analyses will be completed during June of 1971.

IMS Advancement Tests (381-1)

Mr. Ironside

PURPOSE: To develop criterion-referenced test items (grouped into pre- and posttests for each teaching unit) for 10 topics and 10 difficulty levels of the Individualized Mathematics System, a project of RELCV (Durham); items to be based on study of both objectives and instructional materials; items to require manipulatory, perceptual, and/or abstract behaviors by children in grades 1 through 6 or 7. To begin exploration into theory and practice in the development and implementation of such test items.

PROGRESS: ETS Princeton staff developed approximately 400 test items, and also conducted a workshop in Durham for 14 local item-writers who then produced approximately 1200 items under the supervision of SEO staff. All items were reviewed and edited at SEO, and were submitted to the client for the first 5 of the 10 difficulty levels. As agreed, the product was un-pretested raw material, suitable for the client's own pretesting, review, and/or tryout in pilot schools during the period of development. The 1600 test items included about 200 "extension tasks"--an effort to get at the subjects' transfer away from the confines of the particular instructional system. Tests have been printed by the client and have been tried in dozens of pilot classrooms. An informal but final report was submitted in May 1970.

EIP Assessment of Impact (385-1) Mr. Ironside

PURPOSE: To determine the impact of the Durham Education Improvement Program on the local school systems and the community at large over EIP's 5-year tenure in the area. EIP was funded handsomely by Ford Foundation, had as its major goals both experimentation and demonstration in the education of disadvantaged children, and concentrated on intervention programs with children from birth through middle elementary school.

PROGRESS: All phases of investigation have been completed: 90 interviews (with personnel in target schools, LEA administrators, community action groups, parents, other educational agencies, EIP staff); questionnaire survey of school people not directly involved; study of correspondence, official minutes, records, publications, visits. Particular attention was paid to the local and more distant by-products and spin-off effects.

MAJOR FINDINGS: (a) The local school systems have incorporated and internalized virtually nothing of the EIP purposes; (b) local individuals have been much affected by EIP programs and intend to carry reform banners where they can; (c) many persons were unclear as to EIP's numerous segments and their relationships as well as purposes; (d) out-of-town systems were impressed most by the nongraded EIP schools, and several have attempted to emulate this program along with behavior-modification techniques; (e) by-products (vocational training, individualized curricula, school for pregnant girls, distant preschools and ungraded schools) reveal more impact than the intended effects upon the local school systems; (f) in general, positive (and negative) impacts appear to have been transitory.

The report was submitted in August 1970 as PR-70-12.

University of Puerto Rico Interim Project for Physician Performance Criteria (386-5)

Mr. Ironside

PURPOSE: To develop physician performance criteria for use in assessing quality of medicine offered by newly licensed physicians serving in the rural Local Health Centers of Puerto Rico. Secondary purpose was to establish guidelines useful in a number of related educational settings including UPR Medical School M.D. program.

PROGRESS: A committee of 30 UPR faculty and Health Department officials, aided by ETS resource persons, developed, in December 1970, the first run of criteria. Three major areas are hospital practice, ambulatory care, and community medicine. The three subgroups developed quite different styles and degrees of specificity, but came to grips with important dimensions and also many questions related to measurement, relevance, and utility. In their review of the ETS report, the committee is inclined now to feel that it was expecting too much, and is now revising criteria to become "minimal standards" concerned with "any general physician" based mostly on "study of records." A final formal report is expected in summer 1971; it will include (a) the intent and work of the original conference, (b) the more recent and restrictive criteria, and (c) the potential usefulness of the original output.

MAJOR FINDINGS: Community medicine (preventive medicine, public health, education, sanitation, etc.) is apparently easier to deal with in specific performance terms than the other areas. Developing criteria in clinical practice requires the selection of given diseases (or presenting complaints) which in turn calls for selection criteria, not easy to agree on. It is extremely difficult to develop minimum or maximum criteria that are all-encompassing, and easier to work with key indicators in patient management. Nevertheless, the criteria developed here have several important potential (and practical) applications in the Puerto Rican setting.

Advanced Moderator Variable Study
(LSAT Council, 480-15)

Mr. Rock
Mr. Evans

PURPOSE: This study is investigating the moderating effects of certain background variables on the prediction of law school grades. It also is examining whether a number of new tests would act as moderators or predictors.

PROGRESS: Data on the final test battery as well as on the criterion information have been gathered and are now undergoing analysis. Final analysis should be completed by late summer and the final report should be available in fall of 1971.

LSAT Factor Analysis Validity Study (LSAT Council, 480-19) Mr. Carlson

PURPOSE: To gain a better understanding of the specific abilities which contribute to performance on the Law School Admissions Test and to provide suggestions as to the utility of classification of items based on ability as well as the relative predictive validity of the abilities currently being measured. A second aspect of the study was concerned with restructuring and shortening the testing time of each section of the LSAT.

PROGRESS: The following LSAT scores of 941 students entering 18 law schools in 1966 were analyzed: (1) section scores (separately timed sections of the test), (2) factorially defined scores, and (3) selected scores--clusters of items from the Writing Ability test remaining after the morning test factors had been removed and two subsections of Principals and Cases which were not related to any of the factors. In addition, first-year average grades in law school and undergraduate averages were received for each student and standardized for the purpose of combining the data. Correlations and optimal time allocation analyses were carried out.

MAJOR FINDINGS: The validity of factorially designed scores and selected scores is not appreciably greater than that of scores from existing sections of the test. Reallocating testing time among factorially defined scores and selected scores results in a negligible gain over current predictive validity. The testing time for the battery can, however, be reduced to a morning with a significant decrease in predictive validity by omitting the General Background, one of the three Writing Ability sections, and the Figure Classification section.

LSAT Speededness Study (LSAT Council, 480-35)

Mr. Evans
Mr. Reilly

PURPOSE: This study was undertaken to determine if the LSAT is more speeded for candidates who are members of ethnic minorities than for white candidates. Specially constructed "speeded" and "unspeeded" forms of a Reading Comprehension test were administered to samples of regular center and special center LSAT candidates in an effort to determine (1) if the test was more speeded for special candidates and (2) if reducing the amount of speededness was more beneficial to special candidates. Results of the analyses show that (1) the test is somewhat more speeded for special candidates than for regular candidates; (2) reducing the amount of speededness produces higher scores for both regular (22 scaled score points) and special (33 scaled score points) center candidates; and (3) reducing speededness is not more beneficial (in terms of significantly increasing the number of items answered correctly) to special than to regular center candidates. The special centers in the study were at predominately Black colleges. Lower KR-20 reliability was observed under speeded conditions in the special center sample.

PROGRESS: The results of the study are reported in RB-71-31. A replication and extension of this study is currently under way in an attempt to clarify the findings and to gain information about the relationship of validity of speed and power scores and the ethnic background of LSAT candidates.

Discriminant Function Analysis of Law School Performance (LSAT Council, 480-37)

Mr. Schrader
Miss Pitcher

PURPOSE: To test the assumption made in conventional validity studies that the best weights for distinguishing between top students and average students are the same as those for distinguishing between poor students and average students. Data for 26 law schools which participated in the 1968-69 Validity Study Service will be used.

PROGRESS: Analytical work for the first phase of the study, concerned with the hypothesis that more than one canonical function would be needed to describe the three groups, is complete. The second phase of the study designed to compare multiple regression with discriminant function analysis should be completed by November.

MAJOR FINDINGS: In most instances, only one canonical function is statistically significant.

Summary of Validity Studies of Experimental Item Types, 1947-1967 (LSAT Council, 480-42) Mr. Schrader

PURPOSE: For tests designed primarily to predict academic performance, studies of the validity and intercorrelations of various item types provide significant planning data. In the Law School Admission Test a number of validity studies of item types have been made. Although the results of these studies were utilized in making decisions about the test, they are in most instances not readily accessible for consideration in planning new studies. It is planned, therefore, to provide condensed reports of each study, emphasizing the substantive findings.

Development of Undergraduate Rating Scales
(LSAT Council, 480-49)

Mr. Evans
Mr. Hilton

PURPOSE: The purpose of this study is to develop and validate a set of behaviorally anchored rating scales for use as a supplement to or replacement for undergraduate faculty letters of recommendation in the law school admissions process. This project will be carried out over the next three fiscal years. In the first phase of this project law professors, undergraduate faculty, and ETS research staff will cooperate to develop a set of rating scales which will attempt to measure cognitive and noncognitive characteristics of students which are felt to be important to successful law study and at the same time are readily observable at the undergraduate level.

Investigation of Graduate Admissions Policies and Procedures (GRE Board, Phase I, 540-31; Phase II, 540-32) Mr. Burns

PURPOSE: The basic purpose of this project is to collect and disseminate information on graduate admissions policies, procedures, and trends. The objectives will be (1) to provide information for the use of the GRE Board and ETS in planning its programs and activities in graduate admissions; (2) to inform the graduate school community of what kinds of procedures and policies are being followed at various institutions; and (3) to stimulate graduate school thinking and planning about their own admissions procedures and policies. A further objective may be to develop better GRE interpretative information as a result of the information derived from the project.

PROGRESS: The investigation has been conducted in two phases: Phase I involved a survey of all graduate schools, and Phase II consisted of visits of three days' duration to six graduate schools by a member of the ETS staff and a graduate school assistant or associate dean from another school for the purpose of developing a case study of admissions procedures and policies at that institution. The survey and case studies have been conducted. An 85% usable response rate was realized for the survey. A report on the survey entitled Graduate Admissions and Fellowship Selection Policies and Procedures, Part I and Part II, was published in Fall 1970. The case studies were issued as Graduate Admissions and Fellowship Selection Policies and Procedures, Case Studies in Spring 1971. Five Regional Conferences on the general topic were held in Fall 1970.

Alternate Methods of GRE Advanced Tests Mr. Echternacht
(GRE, 540-52)

PURPOSE: The study is to compare the current common-item equating of the GRE Advanced Tests to equating using the Verbal and Quantitative sections of the GRE Aptitude Test. The results will be generalized to the proposed subscores for the GRE Advanced Tests; however, since most of these tests currently have no subscores, the investigation will be made using total scores for most of the tests. The study is designed to answer the following questions: (1) Does equating through the aptitude tests, in general, yield significantly different results than common-item equating? (2) How does the stability of equating through the aptitude tests compare to that of common-item equating across various administration months? (3) How does the stability of equating through aptitude tests compare to that of common-item equating across various educational levels of the candidates?

PROGRESS: All equatings using total sample have been completed.

MAJOR FINDINGS: For five subject areas, score differences between equating methods did not exceed 50 scale points; for eight areas, differences of between 50 and 100 points were found for at least one form; and for four areas, differences of greater than 100 points were found for at least one form. There were no discernible effects resulting from subject area, except for Chemistry, Mathematics, and Physics. Surprisingly, Engineering scores agreed more closely. Also, no differences, resulting from test month or candidates' educational levels, seemed evident, although this portion of the study has not been completed. A report has been presented at the summer 1970 GRE Research and Development Committee Meeting.

The Moderating Effect of TOEFL on GRE	Mr. Sharon
(GRE, TOEFL, 540-62)

PURPOSE: Foreign students desiring to enroll in American graduate schools are frequently examined with both the Test of English as a Foreign Language (TOEFL) and the aptitude tests of the Graduate Record Examinations. No information is presently available as to whether TOEFL adds to the predictive validity of the GRE for foreign candidates. The two tests have been combined through linear regression to predict first-year grade-point average. It was hypothesized that GRE is a valid predictor only for those who score high on TOEFL.

PROGRESS: Twenty-four graduate schools provided GRE and TOEFL scores and grade-point averages for a total of 978 foreign students. The data have been analyzed and a draft report has been submitted to the GRE and TOEFL committees for review.

MAJOR FINDINGS: The hypothesis that English proficiency, as measured by TOEFL, is a moderator in the prediction of graduate grade-point average by the GRE was not supported by the results. The Quantitative GRE test has been found to be the best overall predictor of success. In general, foreign students appear to succeed in American graduate schools in spite of scoring more than one standard deviation below the mean on GRE-Verbal.

TOEFL Language Study (CEEB, 575-11)

Mr. Angoff
Mr. Sharon

PURPOSE: The purpose of this study is to determine whether any differences exist among the English language error patterns of six language groups of college candidates who have taken TOEFL. The results should provide evidence relating to the construct validity of TOEFL and indicate to teachers of English as a Foreign Language the type of instructional emphasis they should give to students of different language backgrounds.

PROGRESS: An item analysis of each of the five parts of TOEFL, separately for each of the six language groups, and an analysis of variance for each part as well as for the total score have been completed. Delta plots, relating each language group to the normative group, have been prepared and measures have been derived expressing the deviation of each item from the central tendency of the items in each plot. A report of the study is now in preparation.

College Environments and Student Academic Achievement (576-3)

Mr. Centra

PURPOSE: The purpose of this study was to investigate selected aspects of the college environment presumed to be associated with student academic achievement in college. Academic achievement was measured by senior students' scores on the Area Tests of the Graduate Record Examination; the Scholastic Aptitude Test (Verbal and Mathematics) scores of these same students prior to college entrance were used as a control measure for differences in initial aptitude. The colleges' social and academic environment were assessed through students' perceptions and included five scales describing the extent of faculty-student interaction, student activism, curriculum flexibility, academic challenge, and the colleges' cultural facilities.

PROGRESS: This study has been completed and a report is available (RB-70-69).

MAJOR FINDINGS: All but the Activism scale were related to student over- or underachievement on one or more of the three Area Tests (Humanities, Natural Science, Social Science). In particular, students at colleges with high scores on the Faculty-Student Interaction scale more often overachieved on two of the criteria tests, while students at colleges with low scores on this scale underachieved on all three of the tests. The results suggest that certain student-described college environmental features are related to academic achievement, although replication with another group of colleges would be desirable.

Disadvantaged Children and Their First School Experiences: A Longitudinal Study (ETS-Head Start, 581-12) Mrs. Shipman

PURPOSE: The purposes of this study are to identify the components of early education that are associated with the cognitive, personal, and social development of children; determine the environmental and background factors which influence such associations; and describe how these influences operate. The children will be followed longitudinally from pre-Head Start through the end of the third grade.

PROGRESS: Field Operations: Given budgetary constraints, data collection efforts were focused on Lee County where there was a relatively large intact sample and few children had had any preschool experience as yet (since Head Start is a kindergarten level program there). As of July 1971, 1188 children were tested individually in three of the four study sites (Lee County, Ala.: 502; Portland, Ore.: 365; Trenton, N. J.: 321). Children tested in Lee County were given a 4 1/2 hour test battery consisting of more than 30 measures. During the same period a 1 1/2 hour interview was obtained with the mother, and an hour was spent observing each mother working together with her child on several interaction tasks. In Portland and Trenton sample children were administered a 1 1/2 hour test battery consisting of nine of the measures used in Lee County.

In the Fall of 1970, the teachers in the sample classrooms (the latter defined as classrooms with 50% or more of study-eligible children) at the four test sites were asked to rate their pupils on a variety of social-personal behaviors. Shortly thereafter, each of the classrooms (91 altogether) were observed in action. These observations were made to obtain information on teacher-pupil and peer interaction and on the approaches individual teachers take in dealing with various classroom situations. Such observations were repeated in the Spring of 1971. In a parallel effort group tests were administered to the pupils in these classrooms. Finally, school and health records of children tested this year are being collected. Teachers and Teacher Aides in Lee County also provided us with additional ratings on those Study children attending Head Start and particular demographic and attitudinal information on themselves.

Princeton Office Activities: Equal emphasis in terms of resource allocation during 1970-71 has been placed on coding, processing, and analysis of data already collected. More specifically, coding, editing and internal analyses of almost all 1968-69 (Year 1) data have been completed. Filing, checking transcribing, scoring, coding and keypunching of 1969-70 (Year 2) data are near completion. The cross-sectional data collected during the Spring of 1970 have been fully processed and the scores have been reported to the schools in the four sites.

Data analyses during the year have involved only Year 1 data and have been mainly concerned with the following areas: (1) comparison of groups of children based on certain status characteristics (e.g., sex, age at time of testing, site, later attendance in Head Start or other preschool programs, etc.); (2) preliminary, internal analyses of instruments and test scores; (3) intradomain analyses; (4) interdomain analyses. The analyses under items (1) and (2) have proceeded relatively smoothly. However, due to limited staff resources we are still in the midst of internal analyses of Year 1 measures. The intra- and interdomain analyses (3 and 4) will therefore be completed during September.

MAJOR FINDINGS: See PR-70-20, "Disadvantaged Children and Their First School Experiences: Preliminary Description of the Initial Sample Prior to School Enrollment," August 1971.

Personality Structure and Development in Headstart Children (OCD, 581-12; 779-1) Mr. Emmerich

PURPOSE: As part of the Longitudinal Study of Disadvantaged Children and Their First School Experiences, this study has the following aims: (1) Evaluation of a multivariate observation procedure for deriving personal-social measures on young children during relatively unstructured class periods; (2) Discovery of the structure (organization) of personal-social behaviors in Headstart children and possible variations in structure associated with the child's sex, age at entry, and first versus second semester enrollment; (3) Examination and interpretation of mean differences on personal-social constructs associated with the child's sex, age at entry, first versus second semester enrollment, and their interactions; (4) Using a personality typology based upon the structural analyses, stability and change in personal-social behaviors during the Headstart year will be evaluated. A model for predicting the qualitative directions of such change will be tested; (5) The classroom personal-social measures will be related to independently assessed measures of cognition, style, and personality in the child; (6) It is hypothesized that parent-child and teacher-child relationships will jointly influence the child's personal-social behaviors in the classroom. These hypotheses will be tested by examining main effects and interactions among maternal and teacher variables with regard to their impact upon the classroom personal-social child measures; (7) The personality types will be conceptualized and tested empirically as moderators of functional relationships between educational experiences (e.g., teacher classroom behaviors) and cognitive development (e.g., test performances at the end of Headstart).

PROGRESS: All Headstart period data have been collected, and those bearing on aims 1-4 have been coded and are being analyzed. A report on aims 1-4 should become available early in 1972.

National Assessment (582)

Miss Anderson
Mrs. Barron

PURPOSE: This is a project of the Education Commission of the States (ECS) directed by the staff of National Assessment of Educational Progress (NAEP). The purpose is to assess the attainment of educational goals in selected subject matter areas by a representative sample of 9-, 13-, 17-year-olds and young adults, and to measure progress over time by periodic reassessments.

PROGRESS: Measures were developed for the first assessment in six areas: writing, literature, science, social studies, music, and art. (Other organizations have developed materials for reading, mathematics, career and occupational development, and citizenship.)

The first assessment of writing, citizenship, and science was completed in 1970-71; literature and reading are being administered in 1971-72, and music and social studies are scheduled for 1972-73.

Writing will be assessed again during 1973-74. In preparation for this second cycle, the statement of objectives developed for Cycle 1 were examined and revised; new exercises are being developed. Pretesting of the materials has been completed and all exercises will undergo review at NAEP sponsored conferences of laymen and subject matter specialists. Development of the assessment materials is being carried out under Mr. Robert Jones' direction.

MAJOR FINDINGS: NAEP has written numerous brochures and articles as well as reports of the results of the 1970-71 assessment. A list of publications can be obtained from the office of the National Assessment Information Service.

Preschool Television for Children (584-18) Mr. Ball

PURPOSE: Children's Television Workshop telecasts Sesame Street, a preschool children's show which is both entertaining and educational. The hope is that preschool children who watch will be the better prepared for school experiences.

PROGRESS: During the period July 1, 1968 to December 31, 1968, ETS has provided consultant services to help CTW develop both behavioral objectives and formative research procedures. In addition, a summative research proposal was written by ETS. Instrument-development proceeded from January through June, and in July 60 children were pretested and posttested using these instruments. Variables measured included knowledge of letters, numbers, forms, body parts and relational terms, and skill at classifying and sorting. A parent-interview was developed and tried out as well.

From September to November 1969, 1,300 children were pretested as part of the major evaluation of Sesame Street. These children, all 3-5 years old, came from Boston, Philadelphia, Durham, Phoenix, and northern rural California. From November 1969 to May 1970 the children were observed viewing Sesame Street and amount of viewing was assessed. The content of the show was analyzed in terms of the show's objectives.

In May 1970 the posttests were administered. During the summer the data were analyzed and on November 4, 1971 the results of the first-year evaluation were presented at a press conference in New York City.

MAJOR FINDINGS: In all, 943 subjects were pretested, posttested, and observed and their parents were given questionnaires to provide background information about them. It was learned that the children who watched Sesame Street most frequently gained most in the areas that were major goals of the show. The relationship of learning to viewing held for all subcategories of children included in the analysis: boys and girls; 3-, 4-, and 5-year-olds; disadvantaged and middle-class children; and children who attended school and those who did not. Of particular interest were extremely high gains observed among a small group of Spanish-speaking children who were included in the study.

Perspectives on Law Enforcement: I. Characteristics of Police Applicants (584-19) Mr. Goldstein

PURPOSE: This study was proposed as a joint venture by the New Jersey Police Training Commission and Educational Testing Service in order to increase our knowledge about the characteristics of police applicants and to provide information which might be utilized by police administrators in the formulation of their policies relating to recruitment, selection, classification, and training. There are three major objectives: (1) to describe a sample of police applicants along a number of relevant psychological dimensions and to compare this group with others wherever possible, (2) to determine whether police applicants differ across communities and whether such differences relate to definable community characteristics, and (3) to determine the similarities and differences between those applicants who pass and those who fail the Civil Service examinations. Corollary to this objective is the goal of determining the characteristics of eligibles if the qualification criteria were modified.

PROGRESS: The first six months of the project were concentrated on the selection and development of instruments to be used in the description of the applicant's characteristics. A battery of five instruments was developed which consists of: (1) a Police Background Information Form, (2) the Personality Inventory, (3) the Police Opinion Questionnaire, (4) the Police Knowledge Tests, and (5) the Discretionary Situations booklet.

Between February 21 and June 27, 1970, a majority of the 1356 applicants who took the civil service examination for patrolman in various locations throughout the State also were administered the special battery. The responses to these five instruments and the civil service examination scores constitute the data analyzed for this study.

MAJOR FINDINGS: Responses to each of the instruments were analyzed for the total group of respondents and the "pass" and "fail" subgroups, i.e., those who achieved a passing score on the civil service examination and those who did not.

The "pass" applicant generally has a higher level of education and a lower rate of unemployment than the one who fails. His religion most probably is Catholic and his race Caucasian. On the other hand, the "fail" candidate more likely was raised in a community of less than 10,000 people and his father was an unskilled or semi-skilled worker.

Pass candidates display a more liberal attitude toward minority group members, particularly Spanish-Americans. These applicants in their responses to the Police Knowledge Test, portray their awareness of the danger present in the policeman's job and their sensitivity to the rights of the individual.

Their responses to the three subsections of the Police Opinion Test show the pass candidates to be generally more rational, reasonable and understanding than those who fail. These characteristics are again displayed in the findings from the test of Discretionary Situations in which the pass group solutions are generally more reasonable and less extreme. Although the findings from the Personality Inventory are more equivocal, they help to differentiate the image of the successful applicant from that of his unsuccessful peer. The former is less likely to avoid dangerous and harmful situations; he is more prone to believe in the basic goodness and honesty of others, and he is more likely to be receptive to other people's problems. The final report for this study is available (PR-71-1).

The Continuing Evaluation of Sesame Street (584-21) Mr. Ball

PURPOSE: Selected subjects from the first-year evaluation of Sesame Street are being followed for another year, some of them into school. A second-year evaluation involving new subjects is also under way. The major purpose of the evaluation is to determine what are the effects of viewing Sesame Street on selected groups of preschool children.

PROGRESS: In October 1970, 302 children who had been included in the first-year evaluation of Sesame Street (in Boston, Durham, and Phoenix) and about 500 new subjects in Los Angeles and Winston-Salem were pretested using the measures developed for the first-year evaluation but revised to incorporate new goals. In May 1971, all of the children were posttested and pretest and posttest parent questionnaires were administered. Between October and May, records of the viewing behavior of the children were obtained, and a content analysis of the year's Sesame Street offerings was performed. Teacher ratings were obtained for the 112 second-year children who entered school in September. The data are currently being keypunched and preliminary data analysis is under way. A report will be prepared during July and August.

Evaluation of Selected Aspects of the CTW Utilization Effort (584-24) Mr. Ball

PURPOSE: To aid CTW in its efforts to increase viewing among inner-city children, a descriptive evaluation is being carried out in the New York metropolitan area of the activities of selected viewing centers there.

PROGRESS: Preliminary contact has been made with the N.Y.C. utilization staff, who will supply the names of viewing centers and their leaders. At the end of June, about 60 children in viewing centers will be tested, along with a small number of children not attending such centers. The instruments to be used will be some of the Sesame Street battery, and a specially designed questionnaire for parents. A diary will be kept of activities in the viewing centers. The children will be posttested at the end of August.

Educational Program Audits (585-27) Mr. Seibel

PURPOSE: Federal legislation authorizing Title VII (Bilingual) and Title VIII (Dropout Prevention) of the Elementary and Secondary Education Act requires that each funded project be subjected to an independent educational program audit. Basically, the audit is an independent (of the project) check on the project's evaluation of its product (students), process (treatment), and management (operation).

PROGRESS: Educational Testing Service is serving as Educational Program Auditor for three Title VIII (Dropout Prevention) Projects during 1970-71:

Project-Job #	Project Title	Location	Audit Project Director
585-27	Potential Dropout Recognition and Prevention Program	Chautauqua, New York	J. Robert Cleary
585-50	Project Emerge	Dayton, Ohio	Dean W. Seibel
692-01	Project Stay	St. Louis, Missouri	Daniel Norton

MAJOR FINDINGS: 1971 interim audit reports for all three projects have been prepared and submitted. Final audit reports for 1969-70 were submitted during the summer of 1970, and final reports for 1971 will be prepared this summer. It is expected that the auditing projects will continue during the 1971-72 academic year.

Chicago Title I Evaluation (585-30) Mr. Marco

PURPOSE: To provide help to the Chicago Public School System on the design and execution of evaluation studies for ESEA Title I projects for 1970-71.

PROGRESS: Reports were issued on 1969-70 child-parent centers and questionnaire results from May 1970 (PR-71-4). The ETS staff also conducted another spring survey of opinions about Title I projects and analyzed the items on the revised Omnibus Inventory, an instrument measuring self-esteem, concern for others, and attitude toward school. General consultation on the use of city-wide test scores and other topics was also provided.

Educational System Information Program--Cincinnati (ESIP, 597-2) Mr. Marco

PURPOSE: To develop a system for processing data obtained by item sampling.

PROGRESS: Data obtained from the Spring 1969 and Spring 1970 test administrations in Cincinnati have been processed. A computer system is now functional. In the first year students in grades 9 and 12 were tested in math and socioeconomic problems, respectively. In the second year students in grade 9 were tested in four areas; estimates of total test parameters were made for ability subgroups and schools as well as for total groups. Final documentation has been completed, and a final report is in preparation.

Taxonomy and Survey Project (597-4) Mr. Trismen

PURPOSE: In September 1968, the U.S. Commissioner of Education and the Council of Chief State School Officers signed an agreement for the development of an evaluation system which will prove a data base for local, state, and federal program decisions. The project is a part of this joint federal/state effort. The Project Descriptor Instrument will be designed to: (1) provide precise descriptions of programs and projects funded by federal sources, (2) compare programs across funding sources, and (3) allow examination of the relationships among components of programs and projects.

PROGRESS: An interim report, issued in September 1969, described the site visits to state departments of education in Georgia, Florida, New York, Michigan, and Pennsylvania and to the U.S.O.E. These visits were for the purpose of examining documents descriptive of federally funded projects to ensure coverage of all important aspects of such projects. The resulting taxonomy is also included in this report.

A Project Descriptor Workbook has now been developed. A field test version of this workbook has been mailed to school districts for completion during the summer. The data from this instrument will be used as a pretest of the ability of the workbook to describe characteristics of programs funded by ESEA Titles I, III, VII, VIII, and the Vocational Education Amendments of 1968.

MAJOR FINDINGS: A final report to USOE has been issued (PR-70-21).

Models of Change and Response Uncertainty (599-3) Mr. Trismen

PURPOSE: To apply certain of the models proposed by J. S. Coleman to data collected during an earlier study on measures of attitude (toward subject matter, course/phase, and marks) and achievement.

PROGRESS: Analyses of items and analyses following a more sophisticated Coleman model have been completed. Additional analyses are planned.

Motivation, Teaching, and Learning in the Classroom Mr. Ball
(599-4)

PURPOSE: A large number of studies have investigated the effectiveness of different methods of teaching. Typically, subjects have been randomly assigned to learn under such conditions as discovery, guided discovery, and rule and example. However, results from these studies have been either contradictory or inconclusive. In a new study undertaken by Sam Ball the motivational styles of the learners were assessed, the hypothesis being that certain teaching methods are differentially useful depending on the learner's motivations. For example, discovery techniques might be optimal for very curious learners but of doubtful value for very anxious learners.

PROGRESS: An attempt was made to analyze the data using a multivariate analysis of covariance technique.

MAJOR FINDINGS: Unfortunately the data were badly skewed. Many subjects under each teaching method were unable to respond correctly to the criterion tests. In general, the expected interactions between motivational style and teaching method were not found. An attempt will be made to replicate the study using a sample of middle-class, academically-able high school students.

Supplementary Analysis of Police Applicant Data (599-8) Mr. Goldstein

PURPOSE: As part of the project entitled "Perspectives on Law Enforcement" (584-19), a battery of five instruments was developed in order to describe some characteristics of police applicants. The data collected with this battery have been analyzed chiefly with chi-square. Supplementary analyses to be done under this project include factor analysis of the Police Knowledge Test, the Police Opinion Questionnaire, and the Discretionary Situations. For each of these analyses, the 10 subscales of the Personality Inventory and Selected Police Background Information items will be used in extension. It is anticipated that this supplementary analysis will be of direct application in future studies of police recruitment and selection, training, and assessment of performance.

PROGRESS: Varimax/Promax factor analyses have been computed separately for the Police Knowledge Test, Discretionary Situations, and the three subtests of the Police Opinion Questionnaire, with the Personality Inventory and selected Background Information items in extension. These are currently being examined for their implications.

Cognitive and Affective Dimensions of Differential Achievement on Open and Closed Book Achievement Tests (599-56) Mr. Marco

PURPOSE: To reanalyze data regarding differences between open- and closed-book tests on selected cognitive and affective dimensions.

PROGRESS: Tests of hypotheses concerning the effect of open-book testing have been made. However, an equating problem has stymied relating differences between test modes to cognitive and affective variables.

Selection of MBA Candidates for Later Job Effectiveness (ATGSB, 690-1) Mr. Harrell

PURPOSE: To seek predictors of success in business that could be used in the admission procedure for Graduate Schools of Business. A large sample of seven classes had taken a series of tests while they were in school.

PROGRESS: A follow-up was completed of the fifth class, five years after graduation. This brought the number of cases in business to 434 for the five classes. Analysis was completed for earnings compared with possible predictors. A new criterion of success, getting into general management, was studied in contrast to being only in a single functional field; possible criteria were tested for those in general management. Articles: "The Personality of High Earning MBA's in Small Business," Personnel Psychology, XXIII, (1970), 369-375; "Personality of High Earning MBA's, Stanford Graduate School of Business Bulletin (Fall, 1970), 10-11; "Earnings of MBA's in the Classes of 1961-65," Graduate School of Business, Stanford University Research Paper No. 6 (May 1971); "Differences between Men in Large and Small Businesses," Personnel Psychology (in press). Grants were received from the Stanford GSB to begin a 10-year follow-up, and from Walter Field to continue the five-year follow-up with the class of 1966. Preparations were made to conduct these follow-ups.

MAJOR FINDINGS: Eighty-six men who were in general management were significantly superior on a number of scales to those in single functional fields such as Marketing or Finance. These scales included those that were found to characterize high earners in both large and small companies, i.e., Ascendance, Personnel Interests, Decisiveness, and Individual Background Survey. Men in General Management were also higher on the Supervisory Qualities Scale, which was not true for high earners.

Criterion Study (ATGSB, 690-5) Mrs. Crooks

PURPOSE: Most evidence indicates that aptitude and achievement measures are the best predictors of academic grades and that the addition of noncognitive measures to the predictor battery produces only moderate increments in the accuracy with which grades can be predicted. If more comprehensive criteria could be developed, however, there is reason to believe that noncognitive data might contribute substantially to prediction. The purpose of this study is to develop criterion measures aimed at qualities other than or in addition to grades which graduate schools of business desire in their graduates and which may be related to later career progress. A variety of predictors, including noncognitive data such as biographical information in addition to measures of aptitude and achievement, could then be studied in terms of their relative contributions to the prediction of each of the several criteria of business school performance.

PROGRESS: Phase I of the study, completed in May 1965, involved the determination of qualities to be evaluated, the development of rating scales anchored by behavioral descriptions, and pretesting. Phase II, completed in 1967, involved the development of a biographical data form, a pilot study in two graduate schools of business of the BDF and the rating scales, analyses of the data, and preparation of a report. Phase III, completed in 1969, involved revision of rating scales and the BDF and a further study of data obtained from 11 graduate schools of business.

MAJOR FINDINGS: In Phases I and II selected rating scales were found to be reliable and to provide information not reflected in grades about students' personal characteristics or qualities. The BDF proved to be a satisfactory instrument for research, with selected data related in significant ways to grades and to ratings. The results in Phase III show that selected biographical data do contribute to prediction not only of ratings of personal qualities deemed important to performance in graduate schools of business but also to prediction of grades and to a combined criterion. The comprehensive criterion of performance composed of some proportion of grades and ratings of personal characteristics appears to be a workable concept and is predicted by a mix of academic and biographical variables. The predictors identified in the preliminary analysis and found to hold up in cross-validation deserve further study. The sample in Phase III was smaller than planned in the design of the study because of the difficulty of securing faculty cooperation in providing the ratings, and the findings must be interpreted and applied with caution. A number of possible outcomes of the study are described in the report, together with implications for future research. Phases I and II are described in RB-70-3 (Thomas L. Hilton, Lorne M. Kendall, and Thomas B. Sprecher). Phase III is described in ATGSB Research Bulletin No. 2 (Lois A. Crooks) and in a Research Supplement to this Bulletin. A proposal has been made for a feasibility study for Phase IV, in which the student sample in Phase II will be followed up five years after graduation.

ATGSB Grade Grouping Study (ATGSB, 690-8) Mr. Boldt

PURPOSE: There is rather widespread dissatisfaction with the grade point average as an evaluation of academic performance. One aspect of this dissatisfaction is that courses are included that require a variety of abilities and achievements to the extent that the grade point average is a conglomerate. The purpose of this study is, on a pilot basis, to detect through factor analytic means groups of courses in which students achieve approximately the same relative standing. If groupings are achieved which eventually prove to have generality across schools, the relationship of grades in a group with test scores will provide a sharper interpretation of the scores.

PROGRESS: The study is reported in RB-70-49.

MAJOR FINDINGS: Only one factor was required to account satisfactorily for the observed course grades. Scores on the factors correlate best (and almost entirely) with ATGSB-Q. The level of the correlation is about the same with ATGSB-Q whether the factor score or the grade point average is used. There was no marked tendency for verbal courses and quantitative courses to be distinguished.

ATGSB New Item Type Research (ATGSB, 690-11) Mr. Carlson

PURPOSE: The study is concerned with investigating the validity of three new item types for predicting success in graduate study in business. The three item types are as follows: (1) Common Word requires an individual to summon up associations for three words and then find the one associated word that is common to all three. (2) For Quantitative Comparison the candidate is required to indicate which of two quantities is greater, or to assert their equality, or to assert the lack of sufficient data to reach a judgment. (3) For Practical Business Judgment the first section is in the form of a reading passage which contains information concerning the immediate objectives of a business executive and the various alternative actions by means of which those objectives might be attained. In the second section, data evaluation, the candidate is asked to classify certain of the facts presented in the problem; and in the third, the data application section, the candidate must answer questions that require him to analyze the alternatives available to achieve the stated objective.

PROGRESS: The three new item types were administered in September 1969 to 836 entering first-year students at six graduate schools of business. Score distributions and correlations among the scores have been run for each of the participating schools. Item analysis was run on all items for the total sample.

The second phase of the study was begun in the Spring of 1970, at which time data-collection rosters were sent to each of the schools in order to gather identification information such as ATGSB test scores, FYA, UGR, and dropout data for each student. Correlations and regression analyses were done.

MAJOR FINDINGS: All three item types, Common Word, Quantitative Comparison, and Practical Business Judgment, differentiated reliably among the candidates and were positively related to academic performance in business school. One of these item types, Common Word, was especially difficult for foreign students, a finding which would cast doubt upon the appropriateness of its use in the ATGSB Program.

The new item types were found to add only minimally to prediction of first year grades when ATGSB scores and undergraduate grade-point average were included in the prediction equation. However, Practical Business Judgment was found to be at least as predictive as the present ATGSB Verbal test, and might, therefore, replace some of the item types presently included in the Verbal test if this were deemed advantageous for other reasons.

In-Basket Study (ATGSB, 690-22) Mrs. Crooks

PURPOSE: One of the problems in graduate business school admissions is that of reaching decisions on which applicants to admit from the marginal candidates, i.e., those candidates whose ATGSB scores, undergraduate records, and other available comparative data place them in the lowest category of possible admits but in larger numbers than places are available. A related problem is selecting candidates who not only will be able to complete the academic program successfully, but will also have the attributes to be effective in their later careers. Follow-up studies of MBA graduates have shown little or no relationship between grades as a measure of performance in business school and later career progress, although admissions criteria used at present do predict grades moderately well. It was proposed that an in-basket test, a simulation of a management job with its related problems, be administered to a sample of entering MBA candidates at several schools. The question to be answered is whether performance on such a test is a differential predictor, not only of grades but of personal characteristics as measured by faculty ratings. There is research evidence that performance on such a test is related to progress in management. This study is being funded by the ATGSB Policy Committee.

PROGRESS: The Consolidated Fund In-Basket Test was administered to 260 entering MBA candidates at five schools early in the fall 1969 term: the University of Michigan, the University of Pittsburgh, Rutgers, Washington University (St. Louis), and the University of Washington. Also administered were a biographical data questionnaire, the Myers-Briggs Type Indicator, and the Ghiselli Self-Description Inventory. Faculty ratings were obtained on seven scales developed for the Criterion Study (ATGSB Research Bulletin No. 2, 1971), and schools were asked to furnish undergraduate average, ATGSB scores, and grade point average at the end of the first year. The in-baskets and questionnaires were scored and the results reported individually to participating students. Analysis of the data was carried out in the fall of 1970 and a final report was submitted to the R & D Committee in June 1971. It will be published in summary form as an ATGSB Brief and in complete form as a DRD Project Report.

MAJOR FINDINGS: Results of this study indicate that the level of prediction of grade average achieved by the usual equation (combination of undergraduate average, ATGSB scores, and an index of quality of undergraduate school) is not improved appreciably by the addition of a measure of in-basket performance. However, where faculty ratings are the criterion, the addition of in-basket performance does increase significantly the level of prediction. Cluster analysis showed that in-basket test scores, faculty ratings, and grades appear to be different measures, although in-basket scores are related to a greater degree to ratings than to grades. It appears from these findings that both ratings and in-basket performance contribute information about a student's attributes and abilities not reflected in grade achievement. Follow-up studies of MBA graduates have disclosed little relationship of academic achievement as measured by grades to later career achievement. This relationship will not be improved unless the criteria for selection and achievement in graduate business school are broadened to reflect the abilities and attributes needed for career success. Recommendations for future research in this area were made.

The Moderating Effect of TOEFL on ATGSB (ATGSB, 690-23) Mr. Sharon

PURPOSE: Foreign students desiring to enroll in American business schools are frequently examined with both the Test of English as a Foreign Language (TOEFL) and the Admissions Test for Graduate Study in Business (ATGSB). No information is presently available as to whether TOEFL adds to the predictive validity of the ATGSB for foreign candidates. The two tests were combined through nonlinear regression to predict first-year grade-point average. It was hypothesized that ATGSB is a valid predictor only for those who score high on TOEFL.

PROGRESS: Seven graduate business schools provided ATGSB and TOEFL scores and grade-point averages for a total of 107 of their foreign students for whom these data are available. The data were analyzed and a final report has been submitted to the ATGSB R&D committee.

MAJOR FINDINGS: The original hypothesis of the study was not put to a test because the student sample was a highly selected one, the average TOEFL score being at the 81st percentile on foreign student norms. The data were pooled across schools for regression analyses by means of the central prediction system. ATGSB was found to be a much more valid predictor of GPA than was TOEFL. The validity of the ATGSB was not significantly improved when this test or its parts were combined with TOEFL in a linear or nonlinear regression.

ATGSB Speededness Study (ATGSB, 690-25)

Mr. Evans
Mr. Reilly

PURPOSE: The purpose of this study is to determine if Quantitative sections of ATGSB are more speeded for minority group candidates than for white candidates. Special forms of ATGSB were developed and administered to random samples of ATGSB candidates in February 1971.

PROGRESS: Data analysis is under way. Comparisons of speeded and unspeeded scores will be made for Blacks, Whites, and Mexican American ATGSB candidates. Reliability of the scores obtained under the various speed conditions will also be investigated.

A Comparison of the Prediction of First-Year Average Grades in Graduate Schools of Business for Interrupted and Uninterrupted Students (ATGSB, 690-30) Miss Pitcher

PURPOSE: This study was undertaken to compare the prediction, from undergraduate record and ATGSB scores, of first-year average grades in graduate schools of business for interrupted and uninterrupted students. An uninterrupted student was defined as one who entered graduate business school within one calendar year following completion of his undergraduate work. An interrupted student is one who waited at least one calendar year after completion of his undergraduate work before beginning graduate work.

PROGRESS: Eleven graduate schools of business which had participated in the ATGSB Validity Study Service in 1967-68 through 1969-70 provided information about interruption for students whose basic validity study data were already available from the Validity Study Service files. Results were analyzed separately by school. A report is being prepared.

MAJOR FINDINGS: In general, the results are consistent with those of past studies of interrupted students. That is, there seems to be a tendency for interrupted students to earn average grades that are higher relative to their measured ability than the average grades earned by uninterrupted students. In spite of the statistically significant results the advantage of the interrupted students is only about a tenth of a letter grade. Length of interruption was found not to be useful as a predictor alone nor in combination with undergraduate record and ATGSB scores. The average length of interruption was about three and a half years.

Indicators of College Quality as Predictors of Success in Graduate Schools of Business (ATGSB, 690-31)

Miss Pitcher
Mr. Schrader

PURPOSE: This study was designed to determine the usefulness of selected measures of college quality for predicting first-year average grades in graduate schools of business. A major objective of the study was to compare the use of college means on ATGSB Verbal, ATGSB Quantitative, and LSAT, Astin's "Estimated Selectivity" score, and average faculty compensation, with ATGSB Total mean which has been used routinely in the ATGSB Validity Study Service studies.

PROGRESS: Six measures of college quality were assembled from existing sources and two intercorrelation tables were run, one based on 626 colleges, the other on more than 3,500 graduate business school students from these colleges. Validity coefficients were computed for each of 14 graduate schools of business, using data from the three years of the ATGSB Validity Study Service. A report is being written.

MAJOR FINDINGS: On the combined sample of students in 14 graduate schools of business, correlation coefficients of ATGSB Total mean with the five other indicators of college quality ranged from .97 with LSAT mean to .72 with average faculty compensation. When college quality indicators were used along with undergraduate record and ATGSB Verbal and Quantitative scores, the multiple correlation coefficients for various combinations of indicators were quite similar, in nearly all instances, to those obtained when only ATGSB Total mean was added to the basic predictors.

A Study of Academic Prediction and Growth in Grades 5 through 12 (719-1)

Mr. Hilton
Mr. Werts

PURPOSE: This is a longitudinal study to determine how academic growth is related to characteristics of the school and to characteristics and experiences of the students.

PROGRESS: The sample of schools, which was selected on the basis of region of the country, size and location of school, and number of graduates who go on to college, consists of 27 public high schools and their junior high and elementary feeder schools and six independent preparatory schools for boys. The School and College Ability Test (SCAT), the Sequential Tests of Educational Progress (STEP), and the Test of General Information (TGI) were administered in the fall of 1961 to about 30,000 students in grades 5, 7, 9, and 11. The SCAT and STEP score distributions were obtained separately for high, middle, and low college-sending schools and for the six independent schools.

College Board Achievement Tests in American History and English Composition and the Preliminary Scholastic Aptitude Test (PSAT) were administered to about 6,000 twelfth-graders in January-February 1963. The seniors also completed a Background and Experience Questionnaire (BEQ). Follow-up information about the high school graduates and the dropouts has been obtained. In the fall of 1963, SCAT, STEP, and TGI were administered to about 33,000 students in grades 7, 9, and 11; in addition, grades 7, 9, and 11 completed a BEQ.

In the winter of 1963, the testing plan of January-February 1963 was repeated for the second group of students to become seniors. In the fall of 1965, grades 9 and 11, the original grades 5 and 7, were tested for the third time. The third wave of senior testing took place in January-February 1967. These seniors were the original grade 7. In the fall of 1967, the original fifth-graders were tested as eleventh-graders. The final senior testing took place in January-February 1969.

Numerous studies have now been completed. An annotated bibliography of Growth Study publications is available (PR-69-11).

Education After High School in 1961 and 1967 (719-1)

Mr. Hilton
Mrs. Bower

PURPOSE: To examine post high school education and occupation for a nationwide sample which graduated in 1967 in comparison with a group which graduated in 1961.

PROGRESS: Data obtained in 1967 about post high school (PHS) education from 7,500 students as part of ETS's Growth Study were compared with similar data collected from 45,000 Project Talent students in 1961. The subjects were divided into four ability groups and then subdivided into four socioeconomic status (SES) subgroups, and a comparison made of PHS educational choices, (four-year college, junior college, and technical-trade schools), within these 16 cells.

MAJOR FINDINGS: Percentage enrollment increased in all categories, particularly junior college, and relatively more low SES students of high ability attended college in 1967 than in 1961. A final report is in preparation.

Home Environment and Educational Development
(719-6)

Mr. Hilton
Mr. Lantz
Miss Patrick

PURPOSE: Recent studies have indicated the importance of process variables in the home as predictors of academic status. This study is an attempt to determine the relationship between students' perceptions of process variables in the home and their academic growth. The measures of academic growth are the STEP, SCAT, and TGI. Students' perceptions are derived from BEQ responses.

PROGRESS: Data relative to students' socioeconomic status and perception of family environment for a group of subjects from the Growth Study population have been examined.

A scale made up of those BEQ items related to students' perceptions of family environment was developed and tested for internal consistency. After some revision, scale scores were computed for a segment of the Growth Study sample. These scores along with mathematics achievement at Grades 5, 7, 9, and 11 and curriculum anticipated at Grade 7 or enrolled in at Grades 9 and 11 are currently being investigated by means of path analytic techniques.

Longitudinal Study of Item Responses (719-9) Mr. Benjamin

PURPOSE: The study employs reading comprehension multiple-choice items in a pretest-posttest design to determine the reliability and validity of the probability of learning as a descriptor of curriculum effectiveness. The probability of learning is estimated four ways, each taking guessing into account in a different manner.

PROGRESS: The study has been completed.

MAJOR FINDINGS: The results indicated acceptable reliability and validity for two of the methods but not for the two based on the standard correction formula or the Horst method of measuring item difficulty. The assumptions of the latter method are questioned by the data. The method that makes a slight correction for guessing is identified as demonstrating the greatest validity and is recommended for use in evaluating an instructional treatment as a next step.

Growth as a Function of the Amount of Training (719-13)

Mr. Hilton
Mr. Berglund

PURPOSE: To study the growth of two variables (STEP Mathematics and SCAT Quantitative) in relation to the amount of training in math given during the time period to be studied, i.e., between grade 5 and grade 11.

PROGRESS: Trend lines based on the mean test performance of students exposed to different amounts of math instruction have been obtained for the total Growth Study sample and separately by sex, by school, and by curriculum. With one major exception (the group of students reporting the most exposure) the trend lines were as predicted. A number of possible explanations for the unexpected performance of the "high exposure" group have been investigated, the most likely appearing to be that the responses of the "high exposure" group are not valid. This hypothesis is being checked by means of a specially constructed response bias scale for the BEQ.

MAJOR FINDINGS: As early as the 5th grade the math achievement of students who select more math courses in high school is superior to the achievement of those who elect fewer courses. Generally, growth in math achievement is positively correlated with exposure although not as strongly as predicted.

Multivariate Change Pilot Study (719-15)

Mr. Beaton
Mr. Hilton

PURPOSE: A technique for longitudinal multivariate analysis is being tried out on a subsample of Growth Study subjects using a selected set of measures and biographical information related to science achievement.

PROGRESS: A computer program for exploring the model has been developed and tested on a number of theories of growth. The model produces parameter estimates, model fits, and estimates of the net effect of variables in complex domains.

MAJOR FINDINGS: The preliminary results indicate a particularly effective way of studying complex growth processes.

The Executive Study (732, All Jobs) Mrs. Crooks

PURPOSE: The Executive Study is a general project initiated in 1955 to encompass basic research on executive performance, selection, and development.

PROGRESS: Early projects and research carried on under this heading (Executive Decision Research (Harbets), Team Coordination, Engineering Study, Executive Position Description Questionnaire, Group Dimensions Questionnaire, Early In-Basket Test Studies) have terminated, but inquiries continue to be received with regard to methodology, findings, instruments used, and research going on elsewhere.

Executive Study Conference: On January 1, 1970, the Executive Study Conference was incorporated by the members and will be administered by its officers and an Executive Committee. ETS has become a regular member. This association of companies holds two conferences a year to discuss research being carried on at the management level in selection, development, and training. The two conferences in 1971 are structured around New Approaches and Techniques in Management Development and the New Work Force: Implications for Management.

Executive Study Conference reports: All reports published of proceedings of conferences prior to December 31, 1969 are available from ETS at $10 per copy. A list of the reports is available on request.

In-Basket Development and Research (735, All Jobs) Mrs. Crooks

PURPOSE: This is a general project to include research and development on the in-basket technique and the application of such instruments for various purposes.

PROGRESS: Consolidated Fund In-Basket Test: This in-basket, in which the setting is a community fund in which the subject assumes the role of the paid director, continues to be used in various training, research, and assessment situations. This is the only in-basket exercise currently available at ETS for purchase and use by outsiders. Scoring must be done by trained scorers. The Public Service Commission of the Canadian government is currently using this in-basket in a research study with their Career Assignment Program. Scoring is being done at ETS. Characteristics of their sample are being studied in relation to other samples from our files, including middle management in business and MBA candidates.

Senior Management In-Basket, Public Service Commission of Canada: An in-basket test at the senior management level is being developed to be used with other measures in assessment programs in departments and agencies of the Canadian government. The programs will attempt to identify present staff members with potential for the senior management level. Interviews with senior managers, time and activity logs kept by their secretaries, and administration of the Executive Position Description Questionnaire to all senior managers provided communalities for a model of the senior management job to be simulated. A committee of senior managers are advising on the project. The in-basket tests will be pilot-tested in the Fall of 1971.

Reference Measures for Cognitive and Personality Factors (ONR, 739-1) Mr. Harman

PURPOSE: Under the sponsorship of the Office of Naval Research, ETS is conducting research to identify and to present measures of well-established factors by updating the previous Kit of Reference Tests for Cognitive Factors and by extending it to the noncognitive domain.

PROGRESS: Work in this first year of the project was directed largely toward literature search and consultation with specialists in the field. The latter objective was met, in part, by a two-day factor analysis conference held at ETS in March 1971. Some 20 prominent contributors in the field of factor analysis provided guidance to the research staff on the current status of reference measures for cognitive and noncognitive factors. The advisory group was chosen to represent the professionals that would be most likely to use such reference measures in the future. The conference was designed in order to give an opportunity for the advisers to help us develop plans that would assure the maximum utility of the resulting reference measures to the profession.

Perceptual Factors in Cognition and Personality
(ONR, 741-10)

Mr. Messick
Mr. French

PURPOSE: This study investigates factor analytically the generality in cognitive and personality domains of the perceptual factors of speed and flexibility of closure.

PROGRESS: Experimental tests of verbal and cognitive closure and flexibility in reasoning and problem solving were developed and included in a test battery with selected personality and reference ability measures. The test battery was administered to more than 500 cadets at the U. S. Naval School of Aviation Medicine, Pensacola, Florida. A factor analysis of intercorrelations among 35 scores on perceptual and cognitive tasks has produced about 14 factors, which were rotated to oblique simple structure. Twenty scores from personality measures were also projected into the space by extension methods. Examination of the results indicated factors interpretable in terms of speed and flexibility of perceptual closure, of verbal closure, and of cognitive closure. Intercorrelations among the primary factors were factor analyzed to produce four second-order factors. A hierarchical transformation was performed to place the first- and second-order factors in the same orthogonal framework and to obtain the loadings of the tests on the second-order factors. The second-order factors appear to reflect verbal comprehension, general reasoning, figural articulation (i.e., facility in structuring and differentiating figural stimuli), and symbolic articulation. A report has been completed and will appear soon as a Research Memorandum.

Accuracy of Estimated True Scores
(ONR, 746-3; 746-4; 772-21)

Mr. Lord
Mrs. Hamilton

PURPOSE: To investigate the accuracy with which the true scores of examinees can be estimated from their observed scores, using a nonlinear regression line.

PROGRESS: A mathematical formulation was developed based on a binomial error model. A computer program has been written to carry out the necessary computations. The method has been applied to several sets of data and a report written.

Disattenuated Partial Correlation (ONR, 746-3; 746-4) Mr. Lord

PURPOSE: To find a significance test for the hypothesis that the correlation between two variables disappears when the true score on a third variable is held constant.

PROGRESS: An "optimal asymptotic test" was found, using a statistical approach due to Neyman.

Efficiency of Estimation When There Is Only One Common Factor (ONR, 746-3; 746-4)

Mr. Lord
Mrs. Wingersky

PURPOSE: The purpose of this project was to discover whether the sampling variance of a test validity coefficient can be reduced if the test items have only one common factor.

PROGRESS: Explicit formulas were derived for the asymptotic sampling variances and covariances of the maximum likelihood estimators for factor analysis parameters in the one-factor case. The effect of the number of variables was explored.

MAJOR FINDINGS: When the one-factor model holds for a set of variables, the covariance between any two variables can be estimated more efficiently than by using the ordinary sample covariance. Nevertheless, estimation of the validity of an unweighted composite is not improved. Results are reported in RB-71-16.

Estimating Item Characteristic Curves (ONR, 746-3; 746-4)

Mr. Lord
Mrs. Wingersky

PURPOSE: To estimate item characteristic curves without assumptions as to their shape or mathematical form. To compare these estimated curves with those found under the common assumption of a generalized logistic model.

PROGRESS: A method was found for estimating i.c.c. from estimated true-score distributions.

MAJOR FINDINGS: Surprising agreement was found with the curves obtained under the generalized logistic assumption. A report was published in the March 1970 Psychometrika.

Estimating Item Parameters (ONR, 746-3; 746-4)

Mr. Lord
Mrs. Wingersky

PURPOSE: To find a practical method to estimate the parameters of the item characteristic curve of each item in a given test.

PROGRESS: A maximum-likelihood method was devised and a computer program written.

MAJOR FINDINGS: A practical computer program for estimating item parameters has been developed and is being made available for general use.

The Self-Scoring Flexilevel Test (ONR, 746-3; 746-4) Mr. Lord

PURPOSE: To develop a test that adjusts the difficulty of the items administered to the ability level of the examinee without the complications usually involved in tailored testing.

PROGRESS: An effective procedure was devised for doing this. Test administration is fairly simple, although the test answer sheet and the directions given to the examinee are more complicated than with an ordinary test. A simple number-right score is satisfactory. A theoretical evaluation of the effectiveness of various flexilevel tests has been carried out.

MAJOR FINDINGS: The flexilevel test is described in RB-70-43, to appear in the Journal of Educational Measurement. The theoretical evaluation (RB-71-6) showed that theoretically, at least, flexilevel tests should provide better measurement than conventional tests whenever the conventional test would need to have an unusually wide spread of item difficulty.

Strong True-Score Theory (ONR, 746-3; 746-4)

Mr. Lord
Mrs. Wingersky
Mrs. Hamilton

PURPOSE: The purpose of this project is to find a mathematical model for the details of the relation between true score and observed score, fitting a wide range of ordinary test data, so that one can estimate or predict in advance the mental test performance of groups of examinees in ways not previously possible.

PROGRESS: A report was published in the 1969 Psychometrika describing a method for estimating true-score distributions, also some empirical results. A computer program for estimating true-score distributions was described in the Fall issue of Educational and Psychological Measurement, 1969 and made available for general use. For example, the program can be used to estimate national norms for a test when only a short form can be administered to the norms sample. A discussion of the problem of estimating true-score distributions was published as a chapter in Random Counts in Biomedical Sciences, 1970, The Pennsylvania State University Press.

Tailored Testing (ONR, 746-3; 746-4) Mr. Lord

PURPOSE: To develop a test theory for "branched" or computer-based testing. To use this theory to evaluate a variety of practical designs for such tests.

PROGRESS: A theory was developed, based on item characteristic curve theory and on relevant practical procedures used in bioassay and elsewhere. A variety of practical procedures were theoretically evaluated.

MAJOR FINDINGS: A report of the theory and of the evaluation of various practical tailored testing procedures appears as Chapter 8 of Computer-Assisted Instruction, Testing and Guidance, Harper and Row, 1970. A report of more recent developments appears in the Spring 1971 Educational and Psychological Measurement. A general overview is given in RM-71-2, accepted for publication in the Journal of the American Statistical Association.

Two-Stage Tests (ONR, 746-3; 746-4)

Mr. Lord
Mrs. Hamilton

PURPOSE: A two-stage test consists of a routing test followed by one of several alternative main tests. The purpose of this study was to find good designs for two-stage tests.

PROGRESS: Some 200 different designs were tried out, using computer methods based on test theory.

MAJOR FINDINGS: When there is no guessing, two-stage testing may be as effective as the best practical tailored testing procedures. The report (RB-69-95) will appear in *Psychometrika*.

On the Sampling Theory of Reliability Estimation Mr. Kristof
(ONR, 746-3; 746-4)

PURPOSE: The paper is intended as a contribution to the sampling theory of reliability estimation when a test has been divided into two, not necessarily parallel, parts.

PROGRESS: A report of the work is contained in RB-69-36 and has been published in the Journal of Mathematical Psychology, 1970, 7, 371-377.

MAJOR FINDINGS: Under normality assumptions, a strict t-test of a point hypothesis about the coefficient alpha parameter is derived. The test is then converted to yield confidence intervals for alpha. These techniques remain applicable even when the initial distributional assumption is considerably relaxed. The methods developed here are complementary to the large sample techniques reported in RB-69-25. Worked examples are appended by way of illustration.

Dimensions of Perceived Item Similarity (USPHS, 752-2) — Mr. Messick

PURPOSE: To ascertain the number and nature of dimensions of perceived similarity among personality statements (in terms of similarities in likelihood of endorsement) in an effort to appraise the relative contributions of perceived content consistencies and perceived desirability consistencies in determining judgments of endorsement frequency.

PROGRESS: Twenty-three MMPI items were selected to span six dimensions previously obtained in analyses of MMPI responses. All possible pairs of these 23 statements were presented to 150 college students, who were asked to rate the similarity of the members of each pair with respect to their probability of endorsement. The 150 x 150 matrix of sums of cross-products among individuals was computed and factored according to the individual differences model of multidimensional scaling. Four dimensions of viewpoint about the similarity of these items appeared, and separate multidimensional spaces were derived for each viewpoint. These perceptual spaces were rotated orthogonally, and the dimensions were interpreted. One important distinction reflected in most of these spaces is the separation of desirable from undesirable item content. A report is in preparation. (Study done in collaboration with Dr. D. N. Jackson of the University of Western Ontario.)

Judgmental Dimensions of Psychopathology (USPHS, 752-2) — Mr. Messick

PURPOSE: To ascertain the number and nature of consistent individual viewpoints about the desirability of MMPI items and to see if these judgmental viewpoints are reflected in response consistencies.

PROGRESS: Ratings of the desirability of all 566 MMPI items were obtained from 150 college students. On a subsequent day, these subjects also responded to the same items under standard MMPI self-description instructions. The 150 x 150 matrix of sums of cross-products among individuals was obtained and factored. Thirteen dimensions of viewpoint emerged. Item loadings for each consistent viewpoint were computed and rotated to orthogonal simple structure. The obtained dimensions reflected differential viewpoints about such characteristics as religious preoccupation, difficulty with control of bodily functions, denial of femininity, oversensitivity, socially deviant attitudes, self-deprecation, listless distractibility, worry, and harm avoidance.

The items marking each viewpoint were then scored as response keys for these subjects and for another sample of 160 college students. Most of the resulting response scales exhibited significant reliabilities, with five scales having reliabilities in the seventies. The two viewpoints that were not interpretable produced reliabilities below .25. The viewpoint scale interpreted as "denial of femininity" correlated with female sex -.72. A report has been completed (RB-70-37) and is scheduled to appear in either the April or June 1972 issue of the Journal of Consulting and Clinical Psychology. (Study done in collaboration with Dr. D. N. Jackson of the University of Western Ontario.)

Leveling-Sharpening as a Cognitive Style
(USPHS, 754-7)

Mr. Damarin
Mr. McKenna
Mr. Messick

PURPOSE: This research has three major purposes and several minor ones. The first purpose is the replication and extension of earlier findings by Gardner, Jackson, and Messick on the cognitive style of leveling-sharpening. The tests which are involved in this research include many psychophysical measures. A second purpose is to discover whether the leveling-sharpening trait or other psychological traits can serve as the parameters of the psychophysical law, as Damarin has suggested they should. Since many leveling-sharpening measures are similar in their formal structure both to measures of figural aftereffects and to measures of retroactive inhibition, a third purpose of the research is to discover whether "levelers" also show larger FAE's and more retroactive inhibition as McKenna has hypothesized.

PROGRESS: Forty-seven objective tests, including putative measures of leveling-sharpening, field independence, category width, neuroticism, figural aftereffects, retroactive inhibition, verbal satiation, rote memory, memory span, meaningful memory, and psychophysical judgment, and a variety of questionnaires, including most of the known response styles, were administered to 147 freshman men and women at the College of William and Mary in Williamsburg, Va.

MAJOR FINDINGS: In an earlier study, now completed, cognitive style measures showed an impressive ability to predict individual differences in the recall of photographs of human faces, thus supporting assertions to this effect by Witkin (cf. RB-63-21).

A later study developed evidence that several different types of leveling-sharpening may exist. One of these is reasonably similar to the dimension that Gardner and his associates interpreted as the tendency to assimilate perceptions to memories.

The same research yielded evidence that individual parameters of the psychophysical function may correlate over different types of judgments, giving some support to Damarin's "trait law" hypothesis. A more refined analysis of this data is now under way.

McKenna's factor analysis of memory tests suggests that there are memory styles associated with a preference for errors of omission vs. commission. A report on this work is available (RB-68-28).

Personality Organization in Cognitive and Social Processes (USPHS, 754-7) Mr. Messick

PURPOSE: This project attempted (a) to clarify in a series of multivariate studies the nature of several cognitive styles and to investigate their interrelations with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; (b) to analyze sex differences in cognitive and personality structure; (c) to assess the stability of cognitive styles and the correlates of reliable changes; and (d) to evaluate the role of cognitive styles in interpersonal and social functioning.

PROGRESS: (a) Categorizing Styles and Cognitive Structure (with Nathan Kogan). This study attempts to clarify two basic questions: (1) Is there a general dimension of breadth of categorization yielding individual consistencies in categorizing behavior across varied stimulus domains and measurement procedures; and (2) what are the cognitive, intellectual, and personality correlates of stylistic consistencies in categorization? Group-administered measures of verbal and quantitative category-width judgments, object sorting, cognitive complexity, preference for complexity, and extremity of judgment were administered to approximately 160 college students, along with several other cognitive, intellectual, and personality variables. Factor analyses were performed separately for male and female subjects. The results indicate that categorizing consistencies exist, but that they are multidimensional and their organization is somewhat different for the two sexes. Females appear to display more generality across categorizing tasks than do males: several dimensions relevant to categorizing emerged for both sexes, but they appeared to be somewhat more differentiated and task specific for males. Personality measures were found to be implicated with categorizing and judgmental consistencies in females, whereas they were more directly associated with intellectual abilities in males. Although the emergence of several categorizing factors in this study indicates the operation of differential stylistic tendencies in this domain, the factors were at the same time found to be correlated with each other, suggesting some generality at a second-order level. Factor analyses of the intercorrelations among the primary factors produced three second-order factors for males and two for females. In both sexes, one of the second-order factors reflected general categorizing consistencies that cut across the stimulus and process distinctions embodied in the various first-order factors.

(b) Cognitive Complexity: Dimensions of Role Constructs (with Nathan Kogan). In conjunction with the study of categorizing styles described above, an attempt was made to assess certain aspects of cognitive complexity using multidimensional scaling techniques. Much recent work on cognitive complexity has used a procedure developed by George Kelly for assessing personal constructs--the Role Construct Repertory Test (Rep Test), in which the subject is presented with several triads of figures significant in interpersonal relations, such as mother, self, and best friend. The subject is then asked to indicate which two members of the triad are most alike and to write down both the basis for their similarity and the reason why the third member is different. The number of different constructs listed has been used as a measure of cognitive complexity.

The 160 subjects in the categorizing study described above were administered a modified form of the Rep Test, and they were also asked to judge on a 10-point scale the degree of similarity or difference between all possible pairs of the 15 role figures. Cross-products matrices for these similarity ratings were computed separately for males and females and were analyzed according to the individual differences model of multidimensional scaling. Six dimensions of viewpoint about stimulus similarity were obtained for the females and seven for the males. The individual scores on these dimensions were then correlated with the cognitive and personality measures included in the categorizing study in an attempt to ascertain some of the properties of individuals holding each viewpoint. In this case, meaningful patterns of significant correlations were obtained between the viewpoint scores and such measures as acquiescence, impulsiveness, rigidity, unconventionality, preference for complexity, category width, and conceptual differentiation.

Two composite stimulus spaces were also derived, one to represent the similarity judgments of the males on the average and one to represent the similarity judgments of the females on the average. For the females, four dimensions were required to represent the average similarity judgments among 15 stimuli, whereas for the males three dimensions appeared adequate. The same three dimensions appeared in both spaces, but the fourth female dimension provided a distinction in sex-role relations (they described similarities between "mother" and "sister" on a separate dimension from similarities between "father" and "brother") that was not emphasized by the males. One report is available and another is in preparation.

(c) Cognitive Interference and Flexible Control. The term "constricted vs. flexible control" refers to consistent modes of reacting to interfering and contradictory cues. This cognitive control has been most frequently assessed in terms of interference scores on the Stroop Color-Word Test. In the present study, a group-administered version of the Stroop Color-Word Interference Test was developed, as well as three other group procedures designed to measure susceptibility to cognitive interference in the face of differing amounts of competing response and at various levels of cognitive involvement. An attempt was made in developing these procedures to vary the ease of separating competing cues in the interference conditions, so as to permit an empirical separation of two related processes thought to be important in withstanding cognitive interference effects: (1) facility in restricting attention to the relevant aspects of the stimulus and the response and (2) facility in actively inhibiting the competing response and other disrupting tendencies. The four tasks were administered to approximately 90 college males, along with the original individually-administered Stroop Color-Word procedure and several other cognitive and personality measures. Both control and interference sections of all five Stroop-type tests were found to correlate highly with each other (median r of .53), but specific processes involved in each task were differentiated in a factor analysis, as were the cognitive and personality correlates of the factors. A factor analysis of intercorrelations among nine first-order factors uncovered two second-order dimensions, which appeared to reflect (1) the ability to withstand disruption in serially organized tasks and (2) the ability to isolate and identify the appropriate aspects of the stimulus and response. A regressed interference score was also obtained for each of the five procedures to reflect that part of the interference performance which was independent of the control performance. These regressed scores possessed substantial reliability, and although they were not related to each other, they exhibited different patterns of significant correlations with personality measures.

(d) Leveling-Sharpening and Assimilation Tendencies (with Fred L. Damarin and Virgil McKenna). The cognitive style of leveling-sharpening has been postulated to account for consistent individual patterns of performance in psychophysical tasks and in the relative differentiation of free association responses. These cognitive consistencies have been attributed to individual differences in the activity of the memory process of assimilation. Levelers assimilate in the sense that they tend to merge perceived objects or stimulus events with similar but not identical objects recalled from previous experience.

In an effort to clarify possible multiple determinants of leveling-sharpening consistencies, the present study attempted to specify the concepts of assimilation and contrast in many different ways as a basis for developing a variety of new experimental tests.

One procedure that was highlighted finds its prototype in the Schematizing Test, which has been used as a criterion measure of leveling-sharpening in other laboratories. The Schematizing Test provides two scores: a lag score measures the subject's tendency to underestimate change in a series of squares that gradually increase in size, and an accuracy score measures his ability to estimate magnitudes accurately within a fixed range of stimulus values.

By also including tasks with different types of stimuli and different judgmental conditions, the present study focused upon the construct validity of alternative interpretations of leveling-sharpening and assimilation-contrast. A battery of several experimental psychophysical judgment tasks was administered to 150 college students, along with measures of field-independence, category width, figural aftereffects, retroactive inhibition, verbal satiation, memory abilities, and various personality characteristics. A factor analysis produced 17 factors, which were rotated to oblique simple structure. An analysis of the intercorrelations among the first-order factors produced four second-order dimensions. One of the factors could very plausibly be interpreted as a measure of assimilation or dedifferentiation of memory traces over time, but the lag score from the Schematizing Test was not related to it. Another factor appeared to represent uncriticalness in judgment, a general tendency to judge similar stimuli as the "same" as opposed to "different," but again the lag score was unrelated. However, the schematizing score was found to be associated with one of the second-order dimensions, which could be interpreted in terms of leveling-sharpening. One report has been completed (RB-68-28) and another is in preparation.

(e) Tolerance for Unrealistic Experiences (with Norman Cliff). This dimension, also called tolerance for instability, refers to the differential willingness to accept perceptions at variance with conventional experience. The concept was introduced originally to account for data linking form-boundedness and form-lability on the Rorschach with responses to unusual kinds of perceptual experiences, such as those arising in apparent movement and aniseikonic distortion.

The present study attempted to appraise the reliability of several measures of apparent movement and aniseikonic distortion, and to evaluate the intercorrelations among them for evidence of stylistic consistencies. Measures of aniseikonic distortion and perceived range of apparent movement were obtained from a sample of 50 undergraduates. Moderate reliabilities were obtained for certain measures of aniseikonic distortion. High reliabilities were achieved for movement threshold and simultaneity threshold scores in the apparent movement procedure, but the two thresholds correlated positively with each other (.46). Correlations between aniseikonic measures and the movement and simultaneity measures were small, but some consistencies appeared. There was a slight tendency for distortion measures to correlate more strongly with movement thresholds than with simultaneity thresholds, and some distortion measures correlated in the same direction with both thresholds. Taken together, these findings suggest that scores for range of apparent movement reflect, in addition to a willingness to perceive the illusion of movement, a possible assimilation phenomenon in which subjects vary in their tendency to stick with a particular judgment (either alternation or movement) once they have made it.

(f) Scanning and Focusing (with Harold Schiffman, Duke University, and Diran Dermen). The dimension of scanning-focusing refers to individual differences in the extent of spontaneous attention deployment. Extensive scanners tend to search and peruse a field of interest, surveying both relevant and irrelevant or miscellaneous properties. The present study attempted to assess scanning and focusing consistencies in a wide variety of tasks in an effort to obtain some convergence of correlations to clarify the interpretation of the construct. These procedures included perceptual judgment tasks (such as size and distance estimation), perceptual speed tasks, and perceptual search tasks (in which the subject must locate in a larger field stimuli of a particular class). Since scanning propensities may also be reflected in the manner in which internal memory fields are surveyed, measures of fluency of ideational production and remoteness of response in word association were also included. A specific attempt was made to differentiate between two possible types of scanning--scanning for signal detection and scanning for information seeking. This was done using perceptual search tasks in which the subject was required to locate stimuli (signals) embedded in meaningfully organized visual fields, e.g., locating faces camouflaged in pictorial scenes. Upon completion of the search task, the stimulus materials were removed, and the subject was then asked specific questions about the content of the scenes. Subjects who incidentally take in information about the field in the process of scanning could thus be differentiated from those whose concern is apparently limited to detecting the signal. Measures were also included for facility in detecting signal stimuli in unorganized fields, such as locating four-letter words in arrays of letters. In view of a theoretical and empirical link between scanning and isolation of affect, several tasks yielding measures of affect expression and control were also included, such as early childhood memories, the Holtzman Inkblot Technique (HIT), and a picture preferences procedure that assesses consistent tendencies to like, dislike, or be indifferent to a wide variety of photographs. These procedures were administered to about 100 male and 120 female college students, along with measures of potentially relevant cognitive styles (particularly field independence and category width), intellectual abilities, and personality.

Factor analyses were performed separately for males and females. Although rotations are not yet complete, some striking relations are discernible.

In males, for example, a factor reflecting incidental knowledge of the content of the field was found to be relatively independent of a factor reflecting facility in locating signal stimuli in unorganized arrays; skill in finding faces hidden in organized fields was associated with both dimensions. In females, a factor emerged linking size estimation with form appropriateness and the absence of shading on the HIT, the latter being a traditional Rorschach indicator of conflict about the acceptance of affect. In both analyses, several cognitive dimensions emerged that displayed substantial correlates with inkblot measures and personality scales.

Another feature of this study was an analysis of errors or "wrong" scores, which was undertaken to investigate the possibility that scanners might be differentiated from nonscanners by the type and number of errors they make. Factor analyses of these "wrong" scores revealed several processes that apparently lead to errors on these tasks. These factors include haste and carelessness, intellectual deficiency, premature perceptual closure, impulsiveness, and susceptibility to distracting and embedding contexts.

(g) Cognitive Styles, Defenses, and Eye Movements (with Lester Luborsky, University of Pennsylvania Medical School, Harold Schiffman, Duke University, and Diran Dermen; partly supported by an NIMH grant to Dr. Luborsky).

Extreme scanning has been found by others to be marginally related to ratings of isolation, projection, and generalized delay on the Rorschach (with isolation possibly related to an "information seeking" type of scanning involved in an obsessive concern with exactness to offset doubt and uncertainty, and with projection possibly related to a "signal detection" type of scanning involved in a paranoid concern with accuracy to offset suspicion and distrust). On the other hand, minimal scanning or focusing has likewise been considered to reflect a concern for accuracy and a preference for viewing the world in a narrow discriminating way. The present study attempts to clarify some of these issues by relating measures of scanning and other cognitive styles (leveling-sharpening and breadth of categorizing) to projective measures of defenses and to characteristic patterns of eye movements in viewing affective and nonaffective stimuli. Heart rate measures of arousal were also obtained under basal control conditions and while viewing the affective and nonaffective pictures.

These procedures have been administered to a small sample of about 40 college students. Several eye-movement scores (such as number of fixations, mean time per fixation, mean track length, and scatter) were obtained for the viewing of each of ten photographs. Heart rate measures were also obtained separately for each photograph presentation. The internal consistency reliabilities of these scores over all ten photographs and over subcategories of photographs reflecting sexual, aggressive, and neutral content have been assessed. Many of the scores proved to be substantially or highly consistent across all of the pictures, but a few were reliable for only one or two of the three subcategories. Such differential reliability was to be expected, since differences in stimulus pattern and meaning should have had more of an impact on some of these scores than on others. The interrelations among eye movement, heart rate, defense, cognitive style, and personality measures have been appraised and a report is in preparation.

(h) Style and Persuasibility (with Fred L. Damarin). This study investigates the interrelations among response styles, cognitive styles relevant to susceptibility to field forces (such as field independence and flexible control), and susceptibility to persuasion and opinion change. A battery of such measures was administered to approximately 90 male college students.

A factor analysis of attitude shift scores, response styles, and selected cognitive and personality measures produced ten factors. Although rotation to oblique simple structure has not yet been completed, preliminary results reveal the presence of four factors clearly related to persuasibility, two of them relatively specific to particular attitude sources and two of them more general in that consistencies are displayed in response to several types of attitude source.

(i) Sex Differences in Cognitive and Personality Structure (with Diran Dermen). This study was designed to bear upon several interrelated questions about the nature and generality of cognitive styles. Of foremost concern were (a) the determination of the number and nature of primary dimensions necessary to account for previously observed stylistic consistencies in cognition; (b) an appraisal of the interrelations of these stylistic dimensions with measures of intellectual ability, attitude and belief structures, affect expression and control, academic and creative achievement, and personality; and (c) an evaluation of sex differences in the psychological organization of cognition and personality.

A battery of experimental measures that required over 20 hours of testing time was administered to more than 700 high school juniors and seniors over a two-month period. The sample included approximately equal numbers of males and females, as well as representatives of college preparatory, commercial, vocational-industrial, and art curricula. High school students were sought for this study because of the extreme difficulty of finding comparable samples of males and females at the college level, i.e., samples where either explicit- or self-selection had not occurred differentially by sex on variables relevant to the sex comparison, not only with respect to obviously relevant variables (like ability patterns and interests) but with respect to more subtle ones as well (like maternal education level). The procedures have been scored not only for the specified dimensions under study, but also for various other response properties (such as number of incorrect or unacceptable answers, number of items omitted, and specific response sets) that would help to characterize the test-taking behavior more completely and might provide important controls for clarifying the basis of some of the obtained relations. Since many of the techniques entailed the evaluation of free responses, the scoring itself was a major undertaking that took more than a year to complete. Reliabilities have been assessed on the total sample and are being appraised separately for male and female subsamples. They have, with few exceptions, proved satisfactory for the total sample. Further analyses are in progress.

(j) The Stability of Cognitive Styles and the Correlates of Change (with Diran Dermen). This study attempts to assess the stability of cognitive style measures over a one-year period and to determine the cognitive and personality correlates of any reliable changes. A sample of 135 high school seniors, who had been extensively tested a year earlier as part of the previous study, were subsequently retested on selected measures of field independence, element articulation, form articulation, conceptual differentiation and compartmentalization, analytic vs. relational categorizing, susceptibility to cognitive interference, and scanning.

Test-retest reliabilities over the one-year interval will be determined, and differences in corresponding means and variances between the two testing sessions will be evaluated. Base-free measures of change will be calculated for each test, in order to partition the difference in true scores between the two testing sessions into two components: one completely dependent upon initial position in the first session and one reflecting that part of the second performance that is completely independent of the previous performance. The reliabilities of these base-free change scores will be evaluated prior to computing their correlations with each other and with the cognitive and personality measures of the first battery.

In addition, a secondary focus of this study dealt with the malleability of categorizing consistencies. Subjects were required to sort the objects in the object-sorting measure of conceptual differentiation first under standard instructions and then in a new and different way. Facility in changing a natural category system could thereby be assessed, as well as the manner in which the change was brought about.

These indices of flexibility and style in changing conceptual categories will be correlated with cognitive style measures from the concurrent battery and with cognitive and personality measures from the previous battery.

A General Method for Analysis of Covariance Structures (NSF, 763-10; 763-11) Mr. Jöreskog

PURPOSE: The purpose of this study is to develop a model and a method so general that the same computational algorithm can be used to analyze many kinds of covariance structures.

PROGRESS: A computer program has been written and tested. The applications performed so far include (a) first- and second-order factor analysis, (b) analysis of sets of congeneric tests, (c) analysis of simplexes and circumplexes, (d) analysis of multimethod-multitrait data, (e) analysis of multitest-multioccasion data, (f) path analysis and linear structural relationships, and (g) covariance structure analysis of multivariate random effects models. A separate version of the computer program includes structures on the mean vectors as well as on variances and covariances.

The general model and the method of parameter estimation are described in RB-69-46. Several examples of the usefulness of the general method are given in RB-69-47. These papers have been published in Biometrika. A paper "Factoring the multitest-multioccasion correlation matrix," presented at the Psychometric Conference in honor of Paul Horst, Seattle, June 22, 1969, and subsequently published in "Current Problems and Techniques in Multivariate Psychology--Proceedings of a Conference Honoring Professor Paul Horst" edited by Clifford E. Lunneborg, deals with applications particularly useful in growth studies. This paper is also available as RB-69-62. Another paper, RB-69-97, which has appeared in Psychometrika, deals with the analysis of sets of congeneric tests. Still another paper, RB-70-42, which appeared in British Journal of Mathematical and Statistical Psychology, is on the estimation and testing of simplex models. The computer programs are described in RB-70-15 and in RB-71-1.

A General Method for Estimating a Linear Structural Equation System (NSF, 763-10; 763-11; 772-68)

Mr. Jöreskog
Miss van Thillo

PURPOSE: The purpose of this study is to develop a general method for estimating the unknown coefficients in a set of linear structural equations. In its most general form the method allows for both errors in equations (residuals, random disturbances) and errors in variables (errors of measurement, observational errors) and yields estimates of the unknown coefficients in the structural equations as well as estimates of the residual variance-covariance matrix and the measurement error variances, provided all these parameters are identified.

PROGRESS: Two computer programs have been written, one for the general case and one for the special case when there are no errors in variables. The methods are described in a paper presented at a "Conference on Structural Equation Models" in Madison, Wisconsin, November 12-16, 1970. The paper will be published in a Conference Volume edited by Professor Arthur S. Goldberger. Both computer programs are currently being revised and extended. One extension allows for multiple congeneric measurements of each true variable.

Retest Changes on Personality Scales (USPHS, 767-2)

Mr. Stricker

PURPOSE: A ubiquitous tendency exists for scores on self-report personality scales to change over time, and in the direction of improved "adjustment," even though no experience intervenes that would be expected to produce such changes. This study seeks to clarify the meaning of these score changes by investigating the nature of individual differences in them. The study will determine the generality of individual differences in score changes on diverse personality scales, as well as the relationships of these individual differences to score changes and initial scores on cognitive and personality measures (not of the self-report type) selected for their relevance in testing hypotheses about the causes of these trends.

PROGRESS: Needed cognitive and personality measures were adapted or developed and administered with the Sixteen Personality Factors Test, the California Psychological Inventory, and the Personality Research Form in a pilot study. The data were analyzed and measures selected for use in the main study. The data for the main study have been collected and the statistical analysis is under way.

Effects of Representative Status and Decision Style on Cooperation in the Prisoner's Dilemma (ARPA, 769-1)

Mr. Kogan
Mrs. Hermann

PURPOSE: The study has two major aims. The first is to examine the effect of being a representative of a group on behavior in the Prisoner's Dilemma (PD). To date in most research on the PD, subjects have acted in their own behalf, the interaction with the opponent being of major concern. What happens, however, if the opponents in the PD must not only be concerned about each other's behavior but also about how well they are representing a reference group? The second aim of the study is to explore the effects of decision style on responses in the PD. By decision style is meant a subject's way of approaching a decision-making task--e.g., his flexibility, his risk-taking dispositions.

PROGRESS: A preliminary report of the study was presented at the 1969 EPA meetings. A final report is currently in preparation.

MAJOR FINDINGS: The results suggest that whether or not a person is acting in his own behalf or representing a group in the PD will affect how he behaves. Representatives as opposed to individuals tend to be more cooperative to start with in the PD, particularly the highly committed representative. A relationship was established between personality and behavior in the PD. This success may stem from using the dyad instead of the individual as the unit of analysis and from examining the effects of personality across trials.

Individual and Group Decision-Making under Conditions of Responsibility for Others (ARPA, 769-1) Mr. Kogan

PURPOSE: In virtually all of the published research on individual and group decision-making, the decision outcomes have impinged only upon the individuals and groups making the decisions. Yet there are frequent occasions in the real world where the outcomes of a decision made by an individual or a group impinges upon others. The major purpose of the present study was to determine whether group-induced risky shifts would decline or disappear when the group's decision affected other people. Subjects selected particular probability-payoff combinations (with real monetary consequences) individually and in groups under conditions where outcomes impinged (a) solely upon the decision-makers and (b) solely upon other individuals.

PROGRESS: A report (RB-69-9) has been issued and submitted for publication. A condensed version has appeared in the Proceedings of the 77th Annual Convention of the American Psychological Association, 1969. A final version has been published in Sociometry, 1971, 34, 198-213.

MAJOR FINDINGS: Under conditions of individual choice, subjects made more conservative decisions when outcomes impinged upon others than when outcomes impinged solely upon oneself. Comparing individual with group decisions, shifts in the direction of greater risk taking were observed irrespective of whether decision outcomes affected solely the decision makers or solely other individuals. The magnitude of the shifts were smaller in the latter case, though not significantly so. Thus, an induced responsibility for persons outside one's group does not seem to eliminate the risky-shift phenomenon in group decisions. (Project conducted in collaboration with Mme. M. Zaleska, Laboratoire de Psychologie Sociale, Sorbonne, Paris.)

Negotiation Behavior under Conditions of Observation and Consultation (ARPA, 769-1) Mr. Kogan

PURPOSE: In an earlier study (RB-67-51), leader-delegate dyads came to joint decisions, then were separated into all-leader and all-delegate negotiating groups with instructions to resolve conflicts of interest and achieve consensus. Differences were observed between the two group types, but these were not of great magnitude. We suggested that the differences might have been attenuated by the fact that leaders and delegates could exert no direct influence upon each other after the dyads were separated. In the present experiment, leaders and delegates negotiated in each other's presence. In half of the groups, leaders negotiated while delegates served as observers and consultants; in the other half, the roles were reversed. Differential negotiation outcomes were expected, given the inequality of power and influence between the leader and delegate roles.

PROGRESS: A report (RB-69-55) has appeared, and has been accepted for publication pending revision.

MAJOR FINDINGS: Predicted differences which failed to appear in the previous study (RB-67-51) were observed in the present research. Delegates negotiating in the presence of their leaders (relative to leaders negotiating in the presence of their delegates) were more likely to deadlock, to seek the opportunity for private consultation with their dyadic partners, and to consume more time both in consultation and actual group discussion. An additional finding of interest was the greater dissatisfaction with group decisions expressed by those in the role of observer than by those in the role of negotiator. (Project conducted in collaboration with Dr. H. Lamm and Miss G. Trommsdorff of the Institute for Social Sciences, University of Mannheim, Germany.)

Pessimism-Optimism and Risk Taking (ARPA, 769-1) Mr. Kogan

PURPOSE: An incidental finding in a recent research article points to a trend toward enhanced pessimism following risky shifts in groups. The present study attempts to replicate and extend the foregoing finding in a systematic experimental manner. Shifts in pessimism-optimism were examined following discussions focused on risk; shifts in risk were examined following discussions focused on pessimism-optimism. Appropriate control groups were employed in both cases.

PROGRESS: A report (RB-69-56) has appeared, and a slightly revised version has been published in the Journal of Personality and Social Psychology, 1970, 15, 366-374.

MAJOR FINDINGS: Asymmetrical effects of risk taking and pessimism-optimism on one another were obtained. Discussions focused on risk led to enhanced risk taking and to increased pessimism concerning the success of the selected alternatives. Discussions focused on pessimism-optimism led to enhanced pessimism, but did not bring about a subsequent increase in preferred level of risk taking. (Project conducted in collaboration with Dr. H. Lamm and Miss G. Trommsdorff of the Institute for Social Sciences, University of Mannheim, Germany.)

Risk Taking in the Context of Intergroup Negotiation (ARPA, 769-1) Mr. Kogan

PURPOSE: This study inquires into the relation of "decision freedom" to the risky shift and other aspects of dyadic negotiation. A negotiator with representative status in his reference group presumably negotiates with stronger commitment--less freedom of independent decision--than a member not elected to represent the group. Accordingly, in comparing the dyadic negotiations of representatives with those of nonrepresentatives, it is predicted that the former will (a) depart less from prior group position; (b) take longer to reach agreement; (c) give more favorable estimates of group's satisfaction with the negotiation outcomes; and (d) be more resistant to the psychological forces underlying the risky shift.

PROGRESS: A report (RB-69-35) has appeared and is now published in the Journal of Experimental Social Psychology, 1970, 6, 351-363.

MAJOR FINDINGS: On the whole, the results of the experiment supported the hypotheses outlined above. (Project conducted in collaboration with Dr. H. Lamm of the Institute for Social Sciences, University of Mannheim, Germany.)

Tests of Cultural Development (770-32) Mr. Diederich

PURPOSE: The purpose of this project is to discover types of difficulties in adult books that baffle and discourage many students at some point near age 14 when they must make a transition from juvenile to adult reading if their interest in reading books is to continue.

PROGRESS: Cooperative Tests and Services had already initiated the development of tests on 20 major literary works that are most widely taught in grades 9-12. The editing, extension, tryout, and statistical analysis of far more questions on these works than will be used in the final forms afford an opportunity to explore in depth the difficulties in a sample of adult books to which the largest number of young people are exposed.

The first 14 of these tests have reached final forms and have been administered in a national sample of schools to establish norms. The last six are scheduled for norming during the next school year.

Bargaining in the Mexican Marketplace (771-2) Miss Keiffer

PURPOSE: To examine the bargaining process in a setting where bargaining occurs as a natural event. Three Mexican marketplaces were chosen as bargaining sites. At each site extensive interviews were conducted with local merchants to find out what local bargaining practices were. Items were bargained for using two general bargaining strategies, interest or disinterest in item, and two concession rates. It was expected that certain of these strategies would lead to no sale on the part of the seller, and that certain other strategies would lead to a lower final price for the buyer.

PROGRESS: Interviews and bargaining were conducted during the Summer of 1969. Analyses of bargaining transactions and content analysis of interview tapes have yet to be done.

Discrimination Experienced by Academic Female Psychologists (771-2) Miss Keiffer

PURPOSE: To document subtle and unsubtle discrimination experienced by women psychologists; to see how many women perceived that they had been discriminated against, or that discrimination existed within academic psychology; to see if discrimination was experienced equally in all areas of psychology. Questionnaires were sent to 314 female psychologists and 51 male psychologists who satisfied the following criteria: membership in one of Divisions 1-8 of APA, Ph.D. from an American university, presently employed in an academic setting, preferably a psychology department, majority of jobs since receiving Ph.D. in academic field, under age 65.

PROGRESS: Questionnaires were sent out and returned. Preliminary analyses reveal that approximately half of both the women and men feel that the standards for admission to graduate school in psychology are different for males and females. Forty per cent of the women have perceived discrimination. Divorced women are much more likely to have perceived discrimination than married women, and single women and widows are least likely to mention discrimination. Women who received their Ph.D. after 1950 are more likely to mention discrimination than those who received their Ph.D. earlier. Over half the women report that their advancement has been slower than men in their specialty. There is no difference between those who received their Ph.D. after 1950 and those who received it earlier, and only small differences between women of the various marital statuses. Half the women report that their advancement in terms of salary increases at their present institution has been worse than that of men. Slightly more women who received their Ph.D. after 1950 report doing worse than men on salary increases than those women who got their Ph.D. earlier, and a higher percentage of divorced women report doing worse than men on salary increases than either single or married women. A more complete data analysis is now under way.

The Effect of a Bargainer's Role on Bargaining Behavior   Miss Keiffer
(771-2)

PURPOSE: Bargainers either represent themselves in a labor-management bargaining game or are the bargaining agents for the fraternity that they are pledging. Fraternities at the University of Pennsylvania have agreed to make participation in a bargaining experiment part of their pledging week activities. It is hypothesized that when bargainers are agents they will drive a harder bargain and end up with a higher final contract than they would if they were bargaining for themselves alone.

PROGRESS: Subjects were run at the University of Pennsylvania. Data now being analyzed. (Study being done in collaboration with Allan Teger of the Psychology Department of the University of Pennsylvania.)

Women's Status and Population Control: A Bibliography
(771-2)

Miss Keiffer
Miss Warren

PURPOSE: To compile a complete bibliography on women's status and how it relates to population control. The bibliography will cover the following categories:

1. General works, historical studies, and status of women.
2. Anthropology, religion, and philosophy.
3. Biology.
4. Psychology.
5. Sociology and sociology of the family.
6. Social commentary.
7. Working women and women and the law.
8. Reproduction and population control.
9. Social, psychological, and economical influences on family size.

PROGRESS: The bibliography has been compiled and is available as RM-70-13.

Study of Objective Tests of Personality in Adults (771-6)

Mr. Damarin

PURPOSE: This study is intended to develop new objective (or performance) tests for R. B. Cattell's personality factors and to determine whether these tests and factors will differentiate among college students and persons with one of two types of personality disorder: alcoholism or psychopathic criminality.

PROGRESS: One hundred sixty-eight subjects including 114 students, 24 alcoholics, and 30 convicts with a diagnosis of psychopathic personality were administered 74 objective and self-report personality measures. The data were subjected to two separate analyses. In the first analysis, 57 chiefly objective test scores and three random variables were intercorrelated and subjected to principal axis factor analysis. Twenty-two factors were rotated graphically to simple structure. In the second study, 17 questionnaire variables, taken chiefly from Cattell's 16 PF, were intercorrelated and factored. Five factors were rotated graphically to simple structure. All of the variables were "extend in" to both factor matrices. In later studies, factor scores were computed and the significance of group differences on these scores was tested by analyses of variance.

MAJOR FINDINGS: A number of Cattell's personality factors including Extroversion and Anxiety were clearly identifiable in one or the other study. The data suggest that objective tests might be used to assess certain factors, such as Extroversion, that were previously defined entirely by self-report measures. There were few objective test loadings on the factor identified as Anxiety, however.

Group differences reaching the .01 level or better were observed on seven out of 22 objective test factors and four out of five questionnaire factors. Alcoholics and criminals were notably lower than students on a factor that might be interpreted as Socioeconomic Status or Cultural Advantage. Associated with low scores were poor performance on tests of ideational fluency; poor performance on tests of the recall of meaningful printed, auditory, or pictorial material, and a lack of speed and skill on psychomotor tasks. There is some reason to believe that this factor may not be general intelligence but instead a differential tendency to accept vs. reject school work and related intellectual disciplines.

Response Styles and Second-Order Factors in the Personality Domain (771-7)

Mr. Stricker

PURPOSE: There have been some suggestions that second-order factors on self-report personality scales may reflect response styles, such as acquiescence and social desirability, rather than such content traits as extroversion, anxiety, and neuroticism. This study is intended to test such response-style hypotheses by examining the relationships between standard measures of these variables and second-order factors on Cattell's Sixteen Personality Factors Questionnaire (16 PF) and Jackson's Personality Research Form (PRF).

PROGRESS: The two inventories and standard measures of response styles have been administered to samples of high school students. The statistical analysis is under way. Four second-order and one third-order factor were extracted from the 16 PF, and six first-order and two second-order factors were extracted from the PRF. These factors have been hand rotated, and their relationships to the response style measures are currently being determined by extending the response style measures into the factor matrices.

Cross-Cultural Study of Group Risk Taking Miss Carlson
(771-16)

PURPOSE: The purpose of this study is to investigate, in African and American cultures, the processes by which groups come to decisions with regard to risk. The study involves a replication of some of the Kogan and Wallach experiments on the "risky shift" (i.e., an increase in risk following group discussion). Cultural factors such as a value placed on willingness to take risks and/or the customary methods of arriving at group decisions may be vital determinants of the results found consistently in Western culture.

PROGRESS: A Research Bulletin (RB-70-59) has been issued. It has also been accepted for publication in a special issue of Journal of Personality and Social Psychology to appear soon.

MAJOR FINDINGS: The data were consistent with the presumed greater tradition-orientation among the Africans. The individual decisions of the African Ss were less risky than those of the Americans. In addition, the risky shift appeared to be less characteristic of the African groups. Questions were raised concerning the generality of the risky shift within Western culture. (This project was conducted in collaboration with Dr. Clive Davis of Syracuse University.)

Labeling and Memory in Children (771-23) Mr. Ward

PURPOSE: Subjects who have and use distinctive labels for the stimuli generally perform better in short-term memory tasks than those who do not possess labels or who have labels but do not use them. The present research is directed toward an understanding of the occasions on which labeling will facilitate recall, and of the processes by which it does so.

PROGRESS: Three studies have been conducted with nursery school children. The basic technique involved presenting stimuli to be remembered with or without overt labeling at the time of presentation, introducing a delay period filled with task-irrelevant verbalization to prevent rehearsal, and then testing the child's recognition of the "target" stimuli in an array.

MAJOR FINDINGS: Overt labeling facilitated recall even in the absence of rehearsal and, it was argued, under conditions in which the labeling effect could not be attributed to the strengthening of a trace in "primary memory."

A report on the first of these studies is available as RB-70-53. A paper based on the three studies was presented at the 1971 meetings of the Society for Research in Child Development, and a related report is in press in Developmental Psychology.

Group Decision-Making in a Minimal Social Situation (771-60)

Mr. Kogan

PURPOSE: The role of group discussion in generating enhanced risk taking in groups remains a controversial theoretical issue. The present study employed electronic semi-automated apparatus which permitted subjects to achieve a group consensus without the need for any verbal communication. The decision material consisted of five different probability-payoff combinations of equal monetary expected value. Subjects experienced 50 individual and 50 group trials with pre-programmed feedback (gain or no gain) following each trial. Comparison of individual and group risk levels is expected to yield information directly relevant to the theoretical issue under test.

PROGRESS: A total of 140 males and females--students at the University of Paris--participated in the experiment. Analysis of data has been completed and a report is in preparation. (Project conducted in collaboration with Roger Lambert of the Laboratoire de Psychologie Sociale, Sorbonne, Paris.)

MAJOR FINDINGS: Subjects manifested greater risk taking under group relative to individual conditions. Various group-decision models have been applied to the data. A model based on majority effects best accounts for the group-decision outcomes.

A Field Study of Nonverbal Creativity (771-61) Mr. Ward

PURPOSE: This study explored creativity in a nonlaboratory setting, in which a heterogeneous self-selected group supplied creative products without awareness that their productions would be used in a psychological investigation. Attention was focused on: (a) a search for relations between creativity and subject variables such as sex and socioeconomic status; and (b) an analysis of the characteristics of products which lead to their being judged as more or less creative.

PROGRESS: In March 1969, a New York radio station held a "Little Green Things" contest; listeners were invited to submit humorous and original little green things, with cash prizes for those judged to be best. These objects were obtained from the station and provided the creative products for this investigation. Individual identification was removed to protect the anonymity of contributors; but sex and address of contributor were recorded, the latter being used to obtain median income, education, etc., for the census tract in which the individual resided. This procedure assumed that individual contributors were to some extent representative of the census tract in which they lived. For a subsample of these products, census data were correlated with originality judgments given by four judges. A further subsample was drawn, and a similar analysis was performed for cross-validation. Judges also rated the products on a variety of apparently simple dimensions--humorousness, complexity, amount of effort expended, etc.--in an effort to determine the characteristics of objects which are judged to be original.

MAJOR FINDINGS: Reliable ratings of originality were obtained from the judges, interjudge agreement tended to be higher for objects created by the sender than for those which were sent in their preexisting natural or manufactured state. For the former, assumed to be the only products which provided a valid index of the creativity of the sender, originality was significantly related to two measures of socioeconomic status. A linear combination of "humorousness" and "amount of effort expended" was effective in predicting originality judgments of objects made by the sender, while "humorousness" and "infrequency" predicted the originality of objects sent without alteration.

Informational Redundancy in Messages as a Function of Stimulus Definition and Complexity (771-89)

Mr. Freedle
Mr. Kingsley

PURPOSE: In referential communication a person tries to map perceived features of his environment into language in such a way that the verbal description efficiently specifies his referent. That is, he constructs messages which contain enough information to be relatively unambiguous but, for reasons of economy, do not contain large amounts of superfluous information. It has been suggested (Olson) that people normally provide a certain amount of redundancy in their messages--they include comments which do not add any new information--but this redundancy is generally held to a low level. However, it is not clear how much redundancy people do normally use nor is it clear what variables affect the amount of redundancy used. The present research is an attempt (a) to assess level of redundancy of messages in a very simple situation which limits the number of stimulus features which might be encoded and makes the encoding of only one feature mandatory, and (b) to vary certain characteristics of the stimulus array and of a subject's perceptual set toward the array in an attempt to find conditions which affect redundancy level of messages. Five-item stimulus arrays were constructed such that the differences on a limited number of dimensions were all perfectly correlated across the five items--making it possible to specify uniquely any item in an array by naming one value on only one dimension. Subjects were asked to write a brief message describing a target item in each array so that a recipient of the message could pick that item out. Experimental treatments varied (a) the complexity of the stimuli (i.e., the number of dimensions on which they varied), and (b) the way stimuli were defined for subjects (i.e., as "bookends," vs. "buildings casting shadows," vs. "geometrical forms"). The latter manipulation was designed to vary the implied dimensionality of the stimuli--the extent to which they were perceived as constituted of many vs. few dimensions--independent of their actual dimensionality. The dependent variable, redundancy, was measured by counting the number of dimensions which subjects encoded in their messages, one dimension being the minimum number necessary (i.e., nonredundancy).

PROGRESS: Messages have been collected in two studies, the second largely a replication and extension of the first. Subjects were college-age adults in both studies, drawn, however, from somewhat different types of college populations. A report is in progress.

Perceptual Errors as a Function of Sex-Role Attitudes
(771-89)

Mr. Baskett
Mrs. Darley
Miss Keiffer
Mr. Kingsley

PURPOSE: That values and attitudes can affect perceptual processes, especially under conditions of tachistoscopic viewing, is well known. The purpose of the present research is to examine this phenomenon in an area of social functioning that (a) tends to be highly emotionally charged, and (b) has been undergoing change in recent years and is now an area of some controversy due to the Women's Liberation Movement, namely, sex roles. The present study involves tachistoscopic presentation of pictures portraying both sex role violations (e.g., a female ship's captain) and non-violations of sex roles to subjects who are required to say what each picture depicts. A paper-and-pencil attitude survey will be administered to these same subjects afterwards to get an assessment of the degree to which they hold conventional attitudes towards sex roles. Predictions for the study are: (1) for pictures showing sex role violations there will be less likelihood of correct identification of the sex of the central figure than for pictures portraying conventionally appropriate role behaviors, and (2) these results will be strongest in subjects whose attitudes toward sex roles are most conventional.

PROGRESS: Stimulus pictures are being drawn and apparatus set up preparatory to running subjects.

Intercultural Attitude Study (772-11) Mr. Gulliksen

PURPOSE: This study is using balanced incomplete block designs, complete paired comparisons, and successive intervals designs, with factor analysis and analysis of variance techniques, in comparing attitudes and preferences of different national groups. Eight different questionnaires relating to work, goals, aims, and activities have been given to respondents in Italy, Poland, Sweden, and the United States (Pennsylvania).

The eight questionnaires used were: (1) Prestige of Occupations, (2) Reasons for Work, (3) Choosing a Job, Part I (Methods), (4) Choosing a Job, Part II (Characteristics), (5) Goals of Life, (6) Means or Rules of Behavior, (7) Aims, and (8) Activities. In addition to the students who answered these eight questionnaires, some of the questionnaires were answered by subjects from military groups, lumbermen, manufacturing workers, and supervisors totalling 3,500 subjects. Scale values were obtained and analysis of variance on votes for items or on category choices. Factor analysis of covariances and varimax rotation with means added proved to be a valuable technique for showing the structure of group attitudes.

PROGRESS: Some concepts which appeared in the first three questionnaires and which were expanded in the latter questionnaires were: the subjects' interest in creativity and variety in work, the wish to dominate others, to achieve friendships, to be considered important; learning on the job, and being told what to do as contrasted with discussion of problems.

MAJOR FINDINGS: Consistency of response was indicated by relatively few total circular triads (TCTs) over $p = .025$. Females showed significantly few TCTs than males at the $F_{.95}$ level. Positive rewards in work were: creativity, self direction, change and variety, learning and advancement, recognition, pay and security. The quality of personal and family life took precedence over an interesting job. Other interests outside of work were still lower in scale values. This study was issued as an Office of Naval Research Technical Report in April 1971.

Learning (772-11) Mr. Gulliksen

PURPOSE: A study of learning curves in cats is under way in an attempt to replicate learning curves in divided brain animals.

PROGRESS: Learning data on normal and split-brain cats has previously been analyzed by a curve fitting program, developed by James Ramsay, which maximizes the likelihood function for the Audley-Jonckheere learning model by the method of conjugate gradients. A learning model earlier proposed by Gulliksen is a special case of the Audley-Jonckheere model. In the current analysis, values were obtained for the parameters of the model: Rho_o, Alpha, Beta, Gamma 1 and Gamma 2. The parameter Rho_o represents the initial probability of a correct response, Alpha represents the amount by which a correct response increments the correct habit strength, and Gamma 1 minus Alpha represents the amount by which a correct response increments the incorrect habit strength. Similarly, Beta represents the amount by which an incorrect response increments the correct habit strength, and Gamma 2 minus Beta represents the amount by which an incorrect response increments the incorrect habit strength.

A plot of the "left vs. right brain" values of Rho_o did not show agreement of right with left brain values. The values of Alpha were generally small, as were the values of Gamma 1 minus Alpha. The values of Beta and Gamma 2 minus Beta did not show agreement of right and left brain values. These parameters varied a great deal. The values of Gamma 2 minus Beta were generally near zero.

The program was revised by James Ramsay to allow the values of Gamma 2 minus Beta and Gamma 1 minus Alpha to become negative. The Princeton cat learning data have been reanalyzed by this new program. Data from Dr. Voneida in Cleveland are being reanalyzed also.

Exploratory Factor Analysis
(772-68)

Mr. Jöreskog
Miss van Thillo

PURPOSE: There are two purposes of this study. One is to evaluate the asymptotic efficiency of the generalized least squares (GLS) method of estimation in factor analysis. The other is to develop a new computational algorithm for exploratory factor analysis.

PROGRESS: The GLS method was investigated in RB-71-26 and shown to yield estimates that are scale-free and asymptotically fully efficient.

A new algorithm that may be used to do exploratory factor analysis by any of the three methods, (i) unweighted least squares (ULS), (ii) generalized least squares (GLS), and (iii) maximum likelihood (ML), is described in RM-71-5. The new algorithm is simpler and faster than Jöreskog's method UMLFA. Particularly, Heywood cases are handled in a simpler and more efficient way.

A Computer Model of Everyday Human Thinking (772-74) Mr. Leaf

PURPOSE: A roughly-outlined model of the processes people use in normal thinking has been formulated from recent research and theorizing. The type of thinking modeled is specifically that used by a person in maintaining a valid picture of the outside world and its relations to him. The model is being written as a computer program (in a list-processing extension of PL-I) to force specification of the implicit or vaguely-indicated details of the description of the model. After being coded, the model will be tested for its effectiveness in handling situations on a representation of recent United States foreign affairs.

PROGRESS: The model has been more completely described in a set of flow charts which are currently being transcribed into coded computer statements. The model represents the environment as a set of concepts interrelated by a set of verb-like relations in concept-relation-concept sentences. It integrates new information into that representation by a two-stage method: (a) The information is evaluated for truthfulness by processes similar to induction and deduction, and (b) the result is screened for psychological dissonance and modified if necessary before being added to the representation of the environment.

A List Processing Extension to PL-I (772-74) Mr. Leaf

PURPOSE: PL-I is a relatively new computer language described by IBM as an extremely sophisticated language combining algebraic power with the ability to deal with data in a wide variety of forms. PL-I is intended to eventually become the standard language for programmers whose applications currently require many different languages. One weakness in PL-I is its poor facility for dealing with alphanumeric or numeric data stored in variable-length lists. Routines are being written which will give programmers easy methods of generating and manipulating such lists while retaining the full power of PL-I.

PROGRESS: The basic routines have been coded in PL-I and 360 assembly language. Patterned after SLIP, the routines produce and manipulate circular lists which can contain data of all basic types allowable in PL-I. Features include the dynamic allocation of space to handle increasing storage requirements of increasing-size lists, simple and informative input/output of list structures, and convenient manipulation of data on the lists.

The routines are debugged for the PL-I version IV compiler and will be stored on OCS.SLIPLIB, which can be used as a linkage-editor automatic call library. Also available is a routine called SNOSCAN, which duplicates the basic SNOBOL text processing facilities for PL-I character strings. A manual which describes the features and uses of PL-I/SLIP and SNOSCAN has been prepared.

Least Squares Matrix Factorization and Multidimensional Scaling (772-75)

Mr. Bechtel

PURPOSE: Although least squares matrix factorization is one of the most widely used psychometric techniques extant, there is no clear and fairly direct proof of this result in psychometric literature. Therefore, the first purpose of the present project is to provide such a proof. The second purpose is to present some theoretical and practical aspects of this result. Finally, the least squares factorization method is applied in several papers in the area of multidimensional scaling.

PROGRESS: A paper on least squares matrix factorization is now complete, except for the section on theoretical and practical aspects. A second paper, which applies matrix factorization to the multidimensional scaling of skew-symmetric (choice) layouts, has appeared as RB-69-73. This was also presented as part of the symposium on "multi-mode factor analysis" at the 1969 Spring meeting of the Psychometric Society. A third paper, which uses matrix factorization in the multidimensional scaling of symmetric layouts has also been written (RB-70-22). A fourth paper uses matrix factorization in a "multidimensional unfolding" analysis developed by Schönemann. This empirical paper involves an unfolding analysis of political choice data which are analyzed by a different model in the second paper. This latter work is currently in the data analysis stage. A fifth paper uses matrix factorization as an adjunct to the analysis of variance in order to least squares fit general nonlinear models to data layouts. A report is available (RB-71-11).

MAJOR FINDINGS: Two major dimensions were extracted from the political choice data by the model presented in the second paper. These dimensions appear to represent "liberalism-conservatism" and "political audacity." On the purely theoretical side, explicit conditions have been found for the exact least squares fitting of nonlinear models to data layouts usually treated by the analysis of variance.

Learning from Verbal Discourse in Educational Media Mr. Carroll
(USOE, 775-1)

PURPOSE: To review the psychological and educational literature on learning from connected meaningful verbal discourse. The theoretical background (or lack thereof) of research in this area will be examined critically as well as the generalizability and applicability of the results to problems of educational media (film, TV, programmed instruction, etc.).

PROGRESS: A 300-page monograph representing a survey of about 1200 references, "Learning from Verbal Discourse in Educational Media," was submitted to the U. S. Office of Education in January 1971. It is understood that this monograph will be published in a collection of other monographs pertaining to various aspects of the psychology of learning from educational media.

A Developmental Study of Attention: A Multivariate Approach (NSF-NICHHD, 776-1)

Mrs. Wilson
Mr. Lewis

PURPOSE: In the study of infant behavior, there is much evidence to indicate that multiple response measures are necessary for a complete understanding of infant development. As Lewis has stressed in recent years, a single response studied at a single age, or even over several ages, can be misleading. Responses change meaning as the infant develops; an infant who cries in response to frustration at two months does not cry in response to frustration at one year; rather, he actively attempts to overcome the frustration. On the other hand, the child who shows little reaction to frustration at two months tends to cry in the face of frustration at a year. Thus, crying changes its meaning--at the early age, it is an active response, but by a year, it has become a passive response. Only a multivariate approach can provide this information.

MAJOR FINDINGS: At least two types of multivariate approaches are possible. The first is the intercorrelational technique in which all responses at each age are correlated with all responses at each other age. A typical result: two measures of attention at different ages, cardiac deceleration at six months and fixation at one year, are strongly related. From this, the investigator infers that at six months cardiac deceleration is the better indicator of attention, but that the two responses change meaning in the second half of the first year, and at one year fixation is the better indicator of attention.

The second approach is the factor analytic technique in which clusters of highly intercorrelated measures are identified. The clusters are then assumed to represent some underlying structure. Thus the clusters, rather than single measures, become the focus of across-age comparisons. One would not expect the clusters themselves to change meaning over age: although the patterning of the responses within a cluster might vary, the underlying structure is fairly stable and the cluster in its service is expected to be fairly consistent over age.

The second approach, using a principal component analysis, was applied to attentional data gathered in a larger longitudinal study. More specifically, age-appropriate pictures were presented to 32 boys and 32 girls, seen longitudinally at 6, 13, 25, and 44 months of age. The child was seated in an infant seat or junior chair in a uniform grey 5' x 5' room. The pictures were presented at eye level by rear screen projection 2 1/2 feet from the child. The measures recorded were first and total fixation time, number of fixations, heart rate, activity, smiling and vocalizing, using the techniques developed by Lewis. The pictures were varied over age in an attempt to maintain interest level. In general, they could be classified as social (having faces), and nonsocial. The pictures were presented in sequence; the time of presentation also varied over age, ranging from 12 to 30 seconds.

A principal component analysis was performed on the data. The number of resulting components, or factors, is determined by the number of measures included in the analysis; if there are six measures, the analysis will compute six components. At all four ages, two components described the data adequately. In all cases, all measures showed significant loadings in at least one of the two components, and the two components together accounted for well over half the total variance. For all ages, the first component accounted for 38 to 41 per cent of the total variance. In each case there were strong positive loadings for first and total fixation and cardiac deceleration, and strong negative loadings for activity. Thus, much looking was associated with little activity. The second component together with the first accounted for 58 to 68 per cent of the total variance for each age. This component was not quite so consistent over the four ages, partially due to the variation in the measures recorded at each age. In general, the component consisted of positive loadings for smiling and vocalizing, and for cardiac deceleration at all but 25 months. In other words, much smiling and vocalizing were associated with much deceleration.

The two components appear to represent quite different patterns of behavior. The first component in general seems to include those responses involved in receptor and body orientation and thus might be said to represent an orienting factor. On the other hand, the second component, though quite variable over age, included an affect measure (smiling or vocalizing) for all but six months. Thus this component might be interpreted as representing an affect factor. These interpretations are supported by further analysis in which a principal component analysis was done on the social and nonsocial data separately, based on the assumption that social stimuli are more apt to produce affect. The results consistently showed the orienting component to be stronger for the nonsocial stimuli (when the affectual responses were at a minimum), and the affect component was consistently stronger for the social stimuli. Thus attention appears to involve at least two types of behaviors --orienting and affect. From this, it is clear that a single response measure, such as fixation, even when measured across age, can tell only part of the story of attentional development. A paper reporting on this study was presented at the Eastern Psychological Association meeting in New York in April 1971.

Developmental Study of Attention within the First Two Years of Life (NSF, 776-1)

Mr. Lewis
Mrs. Wilson
Mrs. Baumel
Mrs. Ban
Miss Van Wyk

PURPOSE: The purpose of the study is six-fold: (1) To study response decrement developmentally across age during the first two years of life. (2) To study cross-modality similarities in response decrement, e.g., visual and auditory modalities. (3) To study the effects of complexity in both the auditory and usual modalities, specifically their effect on response decrement. (4) To compare response decrement to the mother-infant interaction. (5) To compare response decrement to other measures of cognitive skill, specifically the Piaget Scale of Object Permanence and the Bayley Mental Development Scale across the first two years. (6) To compare birth condition to attentional differences.

PROGRESS: A cross-sectional sample and a longitudinal sample of infant subjects are being observed at 3, 6, 9, 12, 18, and 24 months. The sample will be completed in September 1971.

Infants' Attentional Distribution across Two Modalities (NSF-NICHHD, 776-1)

Mrs. Baumel
Mr. Lewis

PURPOSE: This investigation was concerned with two areas which, although important, have been neglected in previous work in infant attention. The first involves attention to repetitive information in several sensory modalities. Much work has been done on response decrement to repeated material in visual, auditory, olfactory and tactile modes. None of the studies, however, have made comparisons between two or more of the modalities. The present study attempts to open up this area of investigation. Responses of three-month-olds to repetitive stimuli in visual and auditory modalities were observed and compared for similarities and differences. The second area investigated in the study was the relationship of the visual and auditory data to measures of cognitive development.

MAJOR FINDINGS: The subjects were 24 three-month-olds, 12 boys and 12 girls. They were seen individually and came to the laboratory on two different occasions. On each visit, the infant was exposed to a visual and an auditory sequence, separated by the administration of one of two cognitive development scales: the Object Permanence Scales, or the Bayley Scales of Infant Development. The auditory stimuli consisted of either a C tone (simple) or a C chord (complex) presented six times with 30-second trials and 30-second intervals between trials. The visual stimuli, three colored lines (simple) or 20 colored lines (complex), were presented for the same number of trials with the same timing. During the sequences several aspects of the infant's behavior were recorded. Fixation (for visual stimuli), cardiac rate, activity, smiling, vocalization, and fretting were the basic measures.

The data indicate several trends within and across modalities. These are summarized as follows. For visual stimuli (measured by fixation time), boys showed discrimination on the basis of complexity while girls did not. For auditory stimuli (measured by cardiac deceleration), neither boys nor girls showed discrimination although there is a tendency towards discrimination in the girls' data. All of the discriminations are in the direction of more attention to complex than to simple events. Total amount of cardiac deceleration indicates a sex x modality difference. Girls' deceleration to auditory stimuli is greater than the decelerations associated with any other events, including boys' and girls' visual and boys' auditory. Response decrement across the six trials was found to be minimal in all areas of measurement. There appear to be two important trends in the data: (1) greater stimulus discrimination in the visual than the auditory mode, and (2) greater cardiac response for girls to auditory than to visual events. However, difficulties arise in comparing the nature of the stimuli in the two modalities. It may be that the simple and complex auditory stimuli are not as distinct on the complexity continuum as the visual stimuli. Thus the greater discrimination found in the visual mode may be a function of inherent stimulus characteristics rather than modality differences. These considerations point to the problems that arise in comparisons across modalities. The finding that girls show greater deceleration in the auditory mode than in the visual mode does not suffer from such theoretical difficulties. Our conclusion is in accord with earlier work (Kagan & Lewis, 1965; Watson, 1969) which suggests that such stimulus-modality differences do exist, with girls responding more to auditory than to visual events. Thus comparisons across modalities are informative, but complex and difficult to interpret. Much more empirical evidence is clearly needed before conclusive inferences can be drawn. A paper, presenting the findings to date, was presented at the Eastern Psychological Association meeting in New York in April 1971.

State as an Infant-Environmental Interaction: Mother-Infant Behavior (NSF - Childhood Council, 776-1)

Mr. Lewis
Mrs. Sarett
Mrs. Watson
Mrs. Wilson

PURPOSE: The present study had three objectives: to define the concept of state, to present a methodology for collecting data on infant-environment interaction--in this case mother-infant--and to demonstrate the former two issues by the presentation of some empirical results.

MAJOR FINDINGS: It is proposed that state be defined in terms of an infant-environment interaction, i.e., state is defined as the interaction of the infant's condition and the nature of the environment. In order to investigate state differences as well as individual differences in state it is necessary to explore and discuss various types of interactive processes and analyses. Empirical data are presented which support the proposed model of state, namely that infant condition is insufficient to describe state since often the same condition has widely different consequences which in turn should affect future infant conditions. The data also reveal individual differences as a function of the sex of the infant. For example, when an infant girl vocalizes, her mother usually vocalizes back; however, when an infant boy vocalizes it is as likely for his mother to touch him as to vocalize to him. These were discussed as an important source of individual variance. A report of the study was presented at the Merrill-Palmer Conference on Research and Teaching of Infant Development, Detroit, Michigan, February 1971 and is available as RB-71-29. It has also been accepted for publication in the Merrill-Palmer Quarterly.

School-Age Children's Comprehension of Words with Multiple Grammatical Functions (USOE, 777-1) Mr. Carroll

PURPOSE: Many words in English have multiple grammatical functions. Sometimes the meaning is essentially the same (e.g., age as a noun and as a verb); sometimes the meaning changes somewhat or even radically from one part of speech to another (e.g., skirt as a noun vs. skirt as a verb). The study attempts to test the hypothesis that children will have more difficulty understanding sentences in which certain words are used in relatively less frequent grammatical functions, as compared with sentences in which these words appear in their more frequent grammatical usages, even when the sense of the word causing the difficulty is related to the sense of the word in its more frequent grammatical function. Studies are to be conducted at the 3rd, 6th, and 9th grade levels using exclusively written material. The study is also considered to have theoretical interest for certain problems of psycholinguistics.

PROGRESS AND MAJOR FINDINGS: The results of this study, completed in January 1971, are reported in RB-71-19, "Comprehension by 3rd, 6th, and 9th graders of words having multiple grammatical functions."

Selective Attention in Children
(USOE, 778-1)

Mr. Hale

PURPOSE: To assess the validity of hypotheses regarding developmental changes in selective attention, children at ages 4, 8 and 12 were given a "component selection" problem adapted from similar paradigms used in research with adults. The task measures the degree to which children attend to component dimensions of stimulus objects (e.g., shape, color, pattern).

PROGRESS: The data collected in the first year of this project have been reported in Research Bulletin 71-3. Data collection for the second major study of this project has recently been completed, and the results are presently being analyzed.

MAJOR FINDINGS: Among the findings reported, for the first major study, a developmental decrease in component selection was observed between ages 4 and 8, with little change from ages 8 to 12. An instructional manipulation designed to increase attention to a single stimulus component had little effect on the children's component selection.

A Longitudinal Study of the Formation of Concepts of Conservation and Classification, as Related to the Development of Spontaneous Cognitive Activities (779-2)

Mr. Wohlwill

PURPOSE: This is a research project currently in progress at Clark University, with support from the National Science Foundation. It aims to assess changes in children's responses to tasks of conservation, class-inclusion and class-intersection over an 18-month period, and to relate these to changes in their approach to situations calling for spontaneous activities involving ordering, measuring and classifying activities.

PROGRESS: An initial test and two retests have been administered to 94 children who were in Kindergarten and Grade 1 at the start of the study. Data for the first two of these sets of tests have been analyzed; analysis of the data for the third test is currently in progress.

MAJOR FINDINGS: On the basis of the data analyzed thus far, suggestive indications have been obtained of a relationship between conservation and measurement, such that a minimal level of proficiency in the latter area is necessary for advances in the former area to occur. Similar relationships have been found between class-intersection and conservation, while a picture of a more nearly parallel progress emerges for the relationship between conservation and seriation activity. Spontaneous classification activities bear no relationship to the acquisition of conservation, and are only slightly related to the formation of concepts of classes.

Methodology and Research Strategy in the Study of Developmental Change (779-2) Mr. Wohlwill

PURPOSE: This project is devoted to a systematic treatment of problems of methodology, research design, data analysis and scientific inference arising in the study of behavioral changes which occur during the course of the individual's development. Particular attention is being given to the following topics: (a) the measurement of changes along behavioral dimensions applicable over extended segments of the age continuum; (b) a discussion of problems of design in developmental research, with particular reference to the uses of longitudinal as compared to cross-sectional data, and variants of these two models; (c) the descriptive analysis of the form and characteristics of the developmental function; (d) the analysis of interrelationships between variables over the course of development; (e) the application of the experimental method to the study of developmental functions; and (f) the integration of the study of developmental changes with the study of individual differences.

PROGRESS: A paper (RB-69-40) discussing the status of the age variable was prepared and is published in Psychological Review, 1970, 77, 49-64. An accompanying paper was distributed as a Research Memorandum (RM-69-29) dealing with several of the above-mentioned problems; it was prepared for a conference on Life-Span Developmental Psychology, held late in April 1969, and is published in P. B. Baltes and L. R. Goulet (Eds.), Theory and Research in Life-Span Developmental Psychology (New York: Academic Press, 1970, Ch. 6). An initial draft of the major portion of a book to be devoted to this topic has been completed.

Developmental Model of Cognitive Style (779-3) Mr. Baker

Study 1: Relations between cognitive controls and school performance in the early grades.

PURPOSE: To identify principles by which children organize and process incoming information from the environment and to relate these individual differences in perceptual-cognitive style to academic success in the early school years. This research is being framed within the context of a particular developmental model, according to which all cognitive controls are not equally available at all stages of life, and a newly formed cognitive control does not replace one that was formed earlier. Rather, the model views cognitive controls as forming a hierarchy of organizations of cognitive behavior, and makes use of the concept of levels within each organization. Focal Attention (i.e., directing attention actively to stimuli and scanning broad sections of the environment so that unique properties of objects are registered) is assumed to emerge first in development. The principle of Field Articulation (i.e., attention directed selectively to relevant stimuli and withheld from irrelevant stimuli guided by the central requirements of the task) emerges next, subordinating and integrating Focal Attention as a distinct process. Leveling-Sharpening (i.e., maintaining in memory the organization of relevant-irrelevant information in the form of a differentiated image with which subsequent information is related and compared) is the third cognitive control principle to emerge and is assumed to subordinate and integrate the preceding two principles. Finally, Equivalence Range is the fourth principle whereby symbols or labels are introduced in order to interrelate bits of information allowing objects to be dealt with in terms of categories or classes. In brief, the present research is aimed at exploring how each of the above described cognitive style dimension relates to school performance in kindergarten and first grade children.

PROGRESS: One hundred fifty white, low to middle class kindergarten children have been given a battery of procedures which assess the above-described cognitive controls, prereading and reading skills, ratings of general classroom behavior (by the teachers), impulse control, and selected personality measures.

It is planned to relate a child's cognitive developmental status, in terms of information processing strategies, with the status of his cognitive skills in terms of reading. The goal is to search for patterns of information processing strategies (and their level of development) associated with the emergence of effective reading skill and to learn more about the relation of certain aspects of personality here presumed to be critical in the learning process and the early formation and operation of cognitive styles.

Fifty first grade children who had been evaluated one year ago were also recently given the same battery of procedures outlined above. The focus here was upon (a) the stability or change in patterns of the cognitive structures and skills under study, and (b) the interrelationship between the assessments of academic success obtained from the teacher and the cognitive status observed during the present testing as well as that observed a year ago.

Data collection for both aspects of this study is now complete, and data analyses should commence shortly. (Study done in collaboration with Sebastiano Santostefano, Boston University School of Medicine.)

Study 2: Individual differences in attentional processes in adults and children.

PURPOSE: Recent research has examined the hypothesis that individuals can be ordered along a dimension in terms of their tendencies to reduce or to augment the intensity of persistent incoming stimulation. At one extreme, the reducer attenuates the intensity of incoming stimulation, thus enabling him to handle high intensities of stimulation quite well but making him acutely uncomfortable when environmental stimulation is minimal (e.g., under sensory deprivation). At the other extreme, the augmenter magnifies the intensity of incoming stimulation, thus showing intolerance of high levels of stimulation but high capacity to cope with very low stimulus intensity. In between is the moderate, who neither greatly enlarges nor minimizes incoming stimulus intensity. Measurement of this dimension has involved variants of the Kinesthetic Figural Aftereffects Task (KFA). Individuals are designated as reducers, as moderates, or as augmenters if a standard block is judged respectively as smaller, as about the same size, or as larger after intervening stimulation. Findings have indicated, for example, that reducers are more tolerant of pain, less able to endure sensory deprivation, more likely to show reduced averaged cortical and visual evoked responses to strong light stimulation, and more likely--if male--to participate in "contact" athletics such as football.

In a series of interrelated studies, the present research seeks to explore the following substantive-theoretical and methodological issues suggested by prior KFA research: (1) Most differential studies utilizing KFA procedure can be criticized either on the basis of the psychophysical procedures employed and/or on psychometric grounds. The first phase of the present research will therefore simultaneously reexamine the reliability and validity of two major variants of the KFA procedure, using normal adults as subjects, and if successful, (2) an attempt will be made to delineate more precisely the nature of the psychological processes underlying individual differences in performance on KFA tasks. (3) A third phase will attempt to delineate educational implications of the augmentation-reduction dimension. The augmentation-reduction typology suggests that the optimum learning environment may be quite different for these three kinds of individuals. The reducer should learn best in an environment in which the techniques that introduce educational materials utilize fairly high levels of sensory stimulation, in which a variety of educational approaches are utilized and in which stimulation is provided by frequent social interaction. However, a calm, nonstimulating environment should best promote learning by the augmenter, who needs far less stimulation than does the reducer and, in fact, would be quite uncomfortable in highly stimulating surroundings. Augmenters should also profit from educational projects which they are given to carry out on their own. (4) Finally, this research will focus upon age changes in the hypothesized augmentation-reduction dimension, and in the correlates of this dimension, with special reference to other possible dimensions which may be tapped by the KFA task.

PROGRESS: Since its inception in the summer of 1970, the following has been achieved: (1) an extensive literature search has been accomplished; (2) a number of substantive and methodological issues which appear to be critical in studying individual differences in kinesthetic figural aftereffects have been tentatively identified, and a research program aimed in exploring them has been broadly sketched; (3) equipment has been secured for administering the two principal variants of the KFA task; (4) an extensive battery of tests and tasks aimed at exploring the first three goals described above has been assembled and/or developed; (5) personnel have been recruited and trained in the administration of the two types of KFA tasks and of the test batteries; (6) five different studies, aimed at exploring the first three goals of the present research delineated earlier, have been concurrently undertaken and have just been completed; (7) based on initial inspection of the data from the first of these studies, a follow-up study with the same subjects is now under way; and (8) preliminary analyses of the data have been undertaken, the data are now in the process of being coded, and more formal data analyses should be under way shortly. (Study done in collaboration with Brian Mishara of Northville State Hospital and Wayne State University.)

The Choice of Prior Probabilities for Bayesian (Logical Probability) Statistical Analysis (NICHHD, 780-2) Mr. Novick

PURPOSE: Bayesian methods of data analysis are becoming increasingly popular in a number of areas of scientific investigation including medical trials and psychometrics. Basic to any method of Bayesian data analysis is the need for the specification of a prior distribution summarizing all of the information available prior to the experiment. This project is concerned with developing methods of specifying reasonable data-oriented prior distributions within the framework of a parametric logical probability model.

PROGRESS AND MAJOR FINDINGS: One method being developed constructs prior distributions using indifference densities (characterizing no prior knowledge) in combination with likelihoods corresponding to hypothetical prior samples that are taken to be approximately equivalent to available prior information. This method was introduced by Novick and Hall and illustrated by Novick and Grizzle in 1965 in papers in the Journal of the American Statistical Association.

Work has been done on applying this method to more complex models than those originally studied. One result was that a satisfactory specification was obtained for the standard mental test theory model when the true- and error-score distributions are taken to be independently and normally distributed. During the second year specifications were obtained for the multivariate normal model and these were related to the normal regression model. During the third year the method was extended to a more general variance components model. These results were presented for discussion before a research methods meeting of the Royal Statistical Society on December 11, 1968 and appeared subsequently in The Journal of the Royal Statistical Society, 1969, 31, 29-64. This model was later developed to cover the problem of regression in m-groups.

Most measurements in medical trials and in psychometric data suffer from errors of measurement. Therefore work on this project has centered on Bayesian analyses for such models. In addition to the normal law test theory model, a model analogous to one studied by Rasch has been investigated. In each case modal estimates of ability parameters have been obtained and in the later case these estimates have been obtained jointly with modal estimates of item difficulty parameters, see RB-69-64. Important social implications of this work to the areas of guidance, central prediction, the choice of predictor variables and multiple comparisons are discussed in RB-69-83. This report later appeared in The Review of Educational Research. Other Research Bulletins issued on this project are 69-31, 69-57, 69-58, 69-54, 69-74, 69-75, 69-91, 70-6, 70-16, 70-32, 70-33, and 71-18. The report 71-18 is the final report on this project and was prepared by the project director at The American College Testing Program. In it a method due to Lindley for empirically generating prior distributions for Bayesian analyses in m-groups is shown, through cross-validation, to be superior to classical methods.

On Proportional Profiles in Factor Analysis Mr. Kristof
(NICHHD, 780-8; 772-63)

PURPOSE: It is intended to examine a number of problems connected with proportional factor profiles. Existence questions, ways of transformation and goodness of fit criteria are studied. Such problems arise, e.g., in multivariate selection situations.

PROGRESS: A number of pertinent theoretical results have been obtained. Their incorporation in a more complete theory is sought.

MAJOR FINDINGS: Necessary and sufficient conditions for the existence of proportional factor profiles on the basis of factor matrices and factor patterns are established. The merits of different goodness of fit criteria are evaluated. The problem of proportional patterns is reduced to the problem of proportional profiles involving factor matrices. The methods employed are based mainly on eigenvectors and eigenvalues.

On a Statistic Arising in Testing Correlation

(NICHHD, 780-8)

Mr. Kristof

PURPOSE: The paper is devoted to the study of a certain statistic, u , defined on samples from a bivariate population with variances σ_{11} , σ_{22} and correlation ρ . If the corresponding statistics are s_{11} , s_{22} and r , then $u = 2r(s_{11}s_{22})^{1/2} (s_{11} + s_{22})^{-1}$. The possible relevance of this statistic in mental test theory is to be explored.

PROGRESS: The work is contained in RB-71-21.

MAJOR FINDINGS: Let the parameter corresponding to u be ν . Under binormality conditions the following is shown. (i) If $\sigma_{11} = \sigma_{22}$, then the distribution of u can be obtained rapidly from the F distribution. Statistical inferences about $\rho = \upsilon$ may be based on F . (ii) In the general case, allowing for $\sigma_{11} \neq \sigma_{22}$, a certain quantity involving u , r and υ follows a t distribution. Statistical inferences about υ may be based on t . (iii) In the general case a quantity t' may be constructed which involves only the statistic u and only the parameter υ . If treated like a t distributed magnitude, t' admits conservative statistical inferences. (iv) The F distributed quantity mentioned in (i) is equivalent to a certain t distributed quantity as follows from an appropriate transformation of the variable. (v) Three test statistics are given, which can be utilized in making statistical inferences about $\rho = \upsilon$ in the case $\sigma_{11} = \sigma_{22}$. A comparison of expected lengths of confidence intervals for ρ obtained from the three test statistics is made. (vi) The use of the formulas derived is illustrated by means of an application to coefficient alpha.

A Theorem on the Trace of Certain Matrix Products and Some Applications (NICHHD, 780-8; 772-63)

Mr. Kristof

PURPOSE: Attainable upper and lower limits of the trace of matrix products of the form $X_1\Gamma_1X_2\Gamma_2\ldots X_n\Gamma_n$ are sought when X_i denotes an orthogonal matrix and Γ_i a diagonal matrix. Matrices X_i are allowed to vary unrestrictedly and independently. Matrices Γ_i are held fixed. Many problems in psychometrics involving the extremization of traces can be solved simply and without the use of awkward and, at times, problematic differentiation when the above-mentioned limits are known.

PROGRESS: A report of the work is contained in RB-69-21, and has been published in the Journal of Mathematical Psychology, 1970, 7, 515-530.

MAJOR FINDINGS: The desired limits of the trace of above matrix products have been established with the help of two lemmas. Complete induction was employed. The usefulness of the theorem is demonstrated by application to a number of examples taken from the literature. Some of them become rather trivial in the presence of the theorem.

On the Theory of a Set of Tests Which Differ Only in Length (NICHHD, 780-8; 772-63)

Mr. Kristof

PURPOSE: A contribution to the sampling theory of a set of homogeneous tests differing only in length is presented, test length being regarded as an essential parameter.

PROGRESS: A report of the work is contained in RB-70-26, which has been accepted for publication in Psychometrika.

MAJOR FINDINGS: Upon formulation of the basic model it is shown that in a combination of such tests forming a "total" test, the signal-to-noise ratio of the components is additive and that the inverse of the variance-covariance matrix of the component measures has all of its off-diagonal elements equal, regardless of distributional assumptions. This fact facilitates the subsequent derivation of a statistical sampling theory. In developing the theory, the cases of known and unknown test lengths are treated separately. For both cases maximum-likelihood estimators of the relevant parameters are derived when the variance-covariance matrix of the components follows a Wishart distribution. Tests of the validity of the model and of hypotheses concerning reliability and standard error of measurement of the total test are also given. It is explicitly shown that in each case the maximum-likelihood equations possess precisely one acceptable solution under rather natural conditions. Application of the methods can be effected without the use of a computer. Two numerical examples are appended by way of illustration.

Extensions of Luce's Choice Theory (NICHHD, 780-25) Mr. Freedle

PURPOSE: This project examines the possibility of modifying Luce's choice models in order to account for asymmetric recognition matrices and asymmetric signal detection ROC curves, for example. His discrimination model has been modified along lines which allow for the effects of "irrelevant" alternatives and perceptual anchors on choice probabilities.

PROGRESS: A report on the sources of asymmetry effects in a syntactic confusion study was presented at the 1969 Psychometric Society meetings and is now available (RB-70-40). A second paper on the effect of irrelevant alternatives on response probabilities in a weight discrimination task (the Shipley-Luce data on two-choice and three-choice weight discriminations) was presented at the 1970 Psychometric Society meetings.

An outline of choice models for discrimination in the presence of irrelevant choice alternatives and for adaptation effects as well as choice models for asymmetric isosensitivity signal detection curves is presented in a review article in press in Psychometrika.

The Psychology of Response Styles (NICHHD, 781-1)

Mr. Damarin
Mr. Messick

PURPOSE: This study aims at learning whether response styles can be explained as personality traits and whether these traits can be measured in situations that do not call for voluntary self-description. Until recently response styles (such as acquiescence, desirable responding and extreme responding) have been conceptualized chiefly as nuisance variables that contaminate tests, especially questionnaires dealing with personality or social attitudes.

PROGRESS: The (large) literature on response styles has been surveyed for evidence linking stylistic responding with personality traits. Correlational and factor analytic studies containing response style measures and objective (or performance) measures of personality have been sought out for evaluation. Twelve published studies by R. B. Cattell provide some evidence that the sought-after relationships can be found. Research Bulletin 65-10 covers this phase of work.

The review suggested a number of ways in which response styles research may be improved. Current measures of the styles themselves may not be optimal. Acquiescence tests, for example, may be contaminated with two other response styles, desirability and inclusiveness (or acceptance), as well as with item content. There are too few conceptually appropriate methodologically sound criterion measures of the personality traits supposedly responsible for stylistic responses. Current theorizing about response styles often seems too narrowly conceived to account for actual laboratory data.

MAJOR FINDINGS: More recent work suggests that response styles in questionnaires may be viewed as facets of the larger problem of self-descriptive behavior. In RB-69-10 Damarin provided a latent structure model in which accuracy of self-description governs the relative prominence of stylistic versus content variance. The model correctly predicted heretofore unobserved effects in the correlation of experimental MMPI scales.

Another version of this model has been successfully applied to personality test responses of preschool children. It correctly predicts the emergence of different response styles at different levels of general intelligence. A report on this work is in preparation.

Creativity and Classroom Behavior (NICHHD, 781-25) Mr. Ward

PURPOSE: An exploratory attempt will be made to determine whether there are identifiable behaviors in the classroom associated with the child's creativity.

PROGRESS: Six classes (fourth through sixth grade) in one urban elementary school have been observed, using a modification of Medley's OScAR procedure which allows identification of the particular child involved in any teacher-child interaction. Creativity and achievement data have also been obtained for the children. Analysis is in progress.

MAJOR FINDINGS: Reliable interjudge agreement was obtained for the major observation categories: number of times pupil initiates an interchange with teacher; teacher initiations which are procedural; teacher's divergent questions; teacher's convergent questions; positive outcomes of interchanges for pupils; and negative outcomes for pupils.

Convergent and Divergent Measurement of Creativity in Children (NICHHD, 781-25) Mr. Ward

PURPOSE: Mednick's conceptualization of the creative process in associative terms has been operationalized in two different ways. This study tested whether, in children, the two kinds of measures do depend on the same processes, and in particular whether individual differences in creativity as measured by Mednick's test depend on the size of the associative repertoire or on the child's ability to evaluate the suitability of possible solutions.

PROGRESS: Fourth-, fifth-, and sixth-grade children were given two forms each of two kinds of creativity tests--divergent measures in which the child named all the ideas he could that met a simple requirement, and convergent measures, adaptations of Mednick's Remote Associates Test, in which he attempted to find one word which was associatively related to each of three others. A report is available (RB-71-40).

MAJOR FINDINGS: As in earlier studies, the divergent measures of creativity were highly intercorrelated and had relatively low correlations with IQ and achievement measures. Divergent and convergent measures shared little variance, and the latter were strongly correlated with IQ and achievement. Moreover, convergent items requiring production of the correct association were strongly related to convergent items requiring only recognition. It was argued that in children Remote Associates performance depends on evaluative abilities rather than on the size of the associative repertoire.

Growth and Stability of Creativity During the School Year Mr. Ward
(NICHHD, 781-25)

PURPOSE: Growth and consistency of individual differences in creativity over a six-month period will be examined as a function of type of creativity test (semantic or figural content), type of score derived (number or mean quality of ideas), ability (standard achievement test scores), and personality (test anxiety and social desirability).

PROGRESS: Data have been collected on six fourth- through sixth-grade classes in one urban elementary school. Analysis is in progress.

MAJOR FINDINGS: Data on number of ideas show highly reliable correlations between the two tests, one with figural and one with semantic content, given in each administration, and similarly high correlations between two forms of either test given six months apart. Number of ideas has significant but substantially lower correlations with ability measures. Quality of ideas shows significant but unimpressive correlations between tasks and over time, and little relation to either number of ideas or ability. Reanalysis of data from three other studies confirms the lack of a coherent dimension of individual differences in idea quality over a broad age and ability range, suggesting that the major source of consistent variance in children's creativity is individual differences in the ability to produce a large number of ideas that are relevant in a simple problem situation.

Some of these findings were presented, in a paper entitled "Creativity Test Performance in Young Children," at the 1971 meetings of the American Educational Research Association.

Cognitive Styles and Psychopathology (NICHHD, 782-3) Mr. Messick

PURPOSE: Recent studies of perceptual and cognitive response patterns of schizophrenics have indicated striking differences between groups of patients with different symptom patterns (paranoid vs. nonparanoid), different levels of adequacy in life adjustment prior to illness (good premorbid vs. poor premorbid adjustment), and different lengths of institutionalization (acute vs. chronic). The present study attempts to utilize the extreme forms of cognitive organization and various personality trait patterns.

PROGRESS: An extensive battery of perceptual, cognitive, and personality measures has been administered to a sample of 95 hospitalized schizophrenics. These procedures were administered either individually or in small groups of four or five patients. The battery included measures of perceptual field-independence (Rod-and-Frame Test, Embedded-Figures Test), element articulation, form articulation, scanning, verbal closure, perceptual closure, extremity and confidence of judgment, category width, risk taking, leveling-sharpening, conceptual differentiation (object sorting), ideational fluency, and verbal comprehension. Also included were items of case history information; a personality inventory assessing several dimensions of neuroses and character disorder; and several rating scales for symptom expression, properties of delusions, and suddenness of onset of illness.

The sample contained roughly equal numbers of paranoids and nonparanoids, who in turn were approximately equally divided into acute and chronic subgroups (using 18 months of hospitalization as the dividing point). The scoring of these measures has been completed, and their reliability for this sample has been appraised and found satisfactory. Various analyses are under way to evaluate subgroup differences on cognitive style and personality measures (in a 2 x 2 breakdown of paranoid vs. nonparanoid and acute vs. chronic patients) and to isolate patterns of cognitive style-personality trait interrelations. (Study done in collaboration with J. Silverman of NIMH and D. N. Jackson of University of Western Ontario.)

Longitudinal Study of Creativity (NICHHD, 782-5) Mr. Kogan

PURPOSE: The long-term stability of measures of creative ability has not been systematically explored. Fifth-grade children who participated in the Pankove-Kogan study (Journal of Personality, 1968, 36, 420-439) have been retested after an interval of five years. Approximately 75% of the original sample was still in the same school systems and hence accessible for retesting. In addition to the readministration of the Wallach-Kogan tasks, the subjects were given a brief questionnaire intended to tap extracurricular accomplishments.

PROGRESS: Analysis of data has been completed, and a preliminary report will be presented at the first symposium of the International Society for the Study of Behavioral Development in Nijmegen, The Netherlands, July 1971.

MAJOR FINDINGS: For boys, correlations between fifth- and tenth-grade composite creativity scores (fluency and uniqueness) were highly significant. For girls, the corresponding correlations were negligible. On the other hand, neither fifth- nor tenth-grade creativity scores were related to extracurricular activities in boys, whereas tenth-grade creativity was marginally associated with such activity in girls. In the latter case, test anxiety operated as a moderator, low anxious girls manifesting a strong positive correlation between tenth-grade fluency and extracurricular activities, in contrast to high anxious girls for whom the correlation was essentially zero.

Differentiation and Development of Social Norms (NICHHD, 782-6) — Mr. Emmerich

PURPOSE: The study examines developmental trends in age-role and sex-role normative differentiations and relates these trends to developmental theories of social learning, cognitive complexity, and cognitive structure.

PROGRESS: Shared beliefs on how the self should act toward others (norms) were studied cross-sectionally in 680 middle-class children and adolescents. Subjects judged the normative import of statements differing in (1) behavioral content (agree with other, help other, seek help from other, argue with other); (2) sex and generation of reference figures holding norms for the self including the self (sources), and (3) sex and generation of recipients of the self's actions (objects).

A follow-up study of certain implications of the findings is now in progress.

MAJOR FINDINGS: The four behavioral contents were differentiated similarly at all ages (continuity), but age-role and sex-role normative differentiations of sources and objects exhibited marked developmental trends (discontinuities), often generalizing across behavioral contents. During middle childhood higher standards were expected in relationships with parents than with peers irrespective of source, whereas during adolescence this differentiation was reversed for attributions to sources within the child's generation. Sex-role norms were strongly sex typed during middle childhood, a pattern that decreased during later childhood and adolescence. Related developmental shifts occurred in choices of identification models, indexed by convergences between norms held by the self and those attributed to other sources. Certain developmental trends were accelerated in brighter subgroups. The study will be published in the Journal of Personality and Social Psychology (with K. S. Goldman and R. E. Shore).

The study provided information on developmental trends in extreme response choice, found to decrease linearly with age in both sexes and earlier (curvilinearly) in girls. Cognitive level was found to mediate choice of an absolute and extreme category, but not choice of a nonabsolute extreme scale category. This aspect of the study will be published as a Brief Report in Developmental Psychology.

In the follow-up study, different aspects of interpersonal relationships are being assessed, e.g., liking A as distinguished from perceiving oneself as similar to A. The study will consider whether different aspects of interpersonal relationships are seen as more differentiated in adolescence than during middle childhood, and whether such developmental trends depend upon whether the object judged is an adult or peer.

Motivation and Capacity in Children's Creativity
(NICHHD, 782-10)

Mr. Ward
Mr. Kogan

PURPOSE: A number of studies have demonstrated that there is a dimension of individual differences in children's creativity which is reliable across a range of alternative tests and which is virtually independent of general intelligence. There is still uncertainty, however, concerning the processes responsible for these individual differences. On the one hand, children who produce more ideas may differ in capacity from those who produce fewer; for example, their "associative hierarchies" may be larger and more richly interconnected. Alternatively, the differences may be found in motivational variables--the productive child may be the one who is more intrigued by an unusual problem, or he may simply have a greater desire to please the examiner.

The effect of a concrete reward on creativity test performance was studied in an attempt to distinguish between these explanations. Reward was intended to assure, so far as is possible, that subjects would approach the task with uniformly high motivation. If the variation in ideational production observed under ordinary testing conditions represents differences in effort expended, reward should eliminate, or at least reduce, the magnitude of those differences. If instead variation in productivity primarily represents capacity, increased motivation should not help the child who lacks relevant associations to draw closer to one with a larger repertoire.

PROGRESS: Four creativity tests modified from those used by Wallach and Kogan were administered to 191 fifth-grade children. In Session 1 all subjects were tested in the evaluation-free, game-like testing context that has been used standardly with these instruments. In Session 2 subjects were assigned to three treatment groups: Control children received a replication of the standard, "baseline" testing conditions. Immediate Reward children were offered one cent for each idea they gave, and were given rewards as they responded. Delayed Reward children were given one cent for each idea, but were not rewarded until the test was completed.

MAJOR FINDINGS: There were no systematic differences associated with the order of testing (tests were counterbalanced to guard against practice effects) or with sex of child. The various creativity tests were highly intercorrelated and were independent of IQ. The main analysis was an analysis of covariance, regressing performance in reward conditions on baseline creativity. Reward had no effect on the slope of the regression, but it led to an increase in mean number of ideas for children at all levels of baseline creativity. It was concluded that a change in the motivational state with which the child approaches the task is not sufficient to eliminate or decrease individual differences in creative performance, and therefore that those differences primarily represent capacity rather than motivational differences among children.

A preliminary report has appeared as a Research Bulletin (RB-70-39) and was presented at the 1970 meetings of the American Psychological Association. A final report is being prepared for publication.

Cognition of Affect (NICHHD, 783-2) Mr. Messick

PURPOSE: This study investigates individual differences in the conceptual differentiation of affects as a function of cognitive styles and ideological commitment.

PROGRESS: Several measures of cognition and personality have been administered to over 600 high school students along with a task in which they were asked to judge the degree of similarity or difference among 18 affect names, such as fear, surprise, joy, and contempt. These judgments of similarity were analyzed by multidimensional scaling procedures to isolate consistent individual viewpoints about affect interrelations. The first two dimensions of judged similarity clearly reflect positive vs. negative affect, and intensity (calm vs. excited). The third dimension is tentatively interpreted as reflective of negative feelings about the self (ashamed, humiliated, afraid) vs. negative feelings directed externally (contemptuous, mad, angry). This last dimension may correspond to intropunitiveness vs. extrapunitiveness. The number of stimulus dimensions and other structural properties of each viewpoint will be used to assess consistent individual differences in the conceptual differentiation of affects. The content of the stimulus representation for each viewpoint will be used to characterize the cognition of affect according to that viewpoint. Individual differences in conceptions of affect will be studied in relation to the cognitive and personality measures included in the battery, but in particular to measures of skill in discriminating and identifying visually displayed affects, the spontaneous tendency to characterize photographs of faces in affect terms, choice of preferred affect, and ideological polarity. The scoring of these procedures is currently in progress. (Study done in collaboration with S. S. Tomkins, City University of New York.)

Dimensions of Conformity (NICHHD, 784-1) Mr. Stricker
Mr. Messick

PURPOSE: This study (a) examines the usefulness of individual differences in ratings of desirability in predicting an independently defined behavioral variable, conformity; (b) investigates the relationship between desirability judgments and self-descriptive responses to parallel forms of personality scales; (c) clarifies the relationship between acquiescence and social desirability (SD) response styles and conformity, a trait which these response styles are believed to reflect; and (d) appraises the dimensionality of responses to diverse conformity situations.

PROGRESS: Pilot study data were used to construct response-style scales based on attitude content and to develop paper-and-pencil and small-group behavioral measures of conformity. Conformity was measured in an Asch situation (counting clicks and answering attitude items) and by questionnaires with fictitious norms (estimating probabilities of events and answering attitude items). The needed personality scales were adapted from existing ones, and all the necessary data were collected from high school student volunteers. Part of the statistical analysis has been completed; the rest is under way.

MAJOR FINDINGS: Many subjects suspected that the purpose of the conformity procedures was to determine whether their responses would be influenced by others, and that the feedback they received about other subjects' responses was wrong. The suspicion variables were highly interrelated and, in general, negatively related to conformity measures. Correlates of the suspicion variables included acquiescence and social desirability response styles, ascendance, self-esteem, and intelligence. A report is published in the Journal of Personality and Social Psychology, 1967, 5, 379-389.

Self-reports and judgments were compared on five personality scales. With a few striking exceptions, the two kinds of measures were generally unrelated, or even negatively related, and they had markedly different patterns of correlations with the conformity measures. These correlations were mediated by the subjects' suspicions about the conformity procedures. Response styles and intelligence did not produce the differential functioning of the two kinds of measures and did not moderate the relationship between them. A report is published in the Journal of Experimental Research in Personality, 1968, 3, 151-167.

The dimensionality of responses to the four conformity procedures and their generality across situations was examined. Conformity, anticonformity, and independence scores on each of these four procedures were factor analyzed, separately for the suspicious and unsuspicious subjects. Conformity and anticonformity appeared to represent one bipolar dimension, and conformity and independence seemed to involve another bipolar dimension. These responses were not consistent over procedures; the devices functioned differently, particularly for the unsuspicious subjects. A report is published in the Journal of Personality and Social Psychology, 1970, 16, 494-507. (This study is being done in collaboration with Dr. Douglas N. Jackson of the University of Western Ontario.)

Dimensions and Categories of Physiognomic Judgment (NICHHD, 784-2) Mr. Kogan

PURPOSE: The general purpose of the present study is to examine the manner in which the use of an age judgment scale varies with the chronological age of the judge.

PROGRESS: Photos of stimulus persons varying in age were presented to samples of college students and elderly people for judgments of chronological age. Both male and female stimulus persons and subjects were employed. Analyses of variance were carried out separately for male and female photos. A report based on this research is in preparation.

MAJOR FINDINGS: A main effect for judge's age was obtained only in the case of the male photos--older judges attributing higher age levels to the stimulus persons. For both male and female photos, a highly significant interaction (age of judge by age of stimulus person) was obtained. Stimulus persons at the younger end of the age continuum were judged consistently older by the older judges. The inflection point is approximately age 40 for males and age 30 for females. Stimulus persons exceeding those ages are judged differently depending upon whether they are male or female. Photos of older females are judged consistently younger by the elderly subjects, whereas the age of the judge does not seem to have a systematic effect on the ages attributed to the photos of older males.

Determinants of Risk Taking for Others (NICHHD, 784-4)

Mr. Kogan
Mr. Teger

PURPOSE: The present research represents an extension of earlier work by Zaleska and Kogan (RB-69-9) on the comparison of individual and group decision-making for other people. The study examines the effects of friendship and reciprocity of decision-making between self and other on preferred risk levels.

PROGRESS: Data analysis has been completed, and a preliminary report will appear in the proceedings of the 1971 APA Meetings.

MAJOR FINDINGS: Decision-making for others was not affected by extent of friendship between the decision-maker and the other. On the other hand, a significant reciprocity effect was obtained. In deciding for another, subjects were inclined to take greater risk under nonreciprocal conditions than in the case where the other would in turn be deciding for the decision-maker herself.

Motivation and Performance in Adult-Evaluative and Peer-Competitive Contexts (NICHHD, 784-4) Mr. Kogan

PURPOSE: In the last 15 years there has been a growing body of research into motivational factors in achievement in both the classroom and other situations. Such research has been possible because of the development of useful measures of achievement motivation. Although psychologists have long recognized the influence of the interpersonal context upon the individual's striving and performance, there has been little research on the specific effects of such contexts. The principal aim of this study is to examine the impact of two commonly occurring contexts--adult-evaluative and peer-competitive--on the achievement motivation and performance of elementary school boys.

PROGRESS: A total of 370 fifth- and sixth-grade boys in the Princeton School System have been tested for need Achievement, need Affiliation, test anxiety, and defensiveness. Of that number, 168 boys varying in "hope for success" and "fear of failure" have performed on a shuffleboard risk-taking task. Analysis of the data has been completed and a report is currently in preparation.

MAJOR FINDINGS: On the whole, the results in both the peer-competitive and adult-evaluative contexts conformed to Atkinson's motivational model of risk taking--i.e., a preference for intermediate risk taking in "hope for success" subjects and a preference for more extreme risk or caution in "fear of failure" subjects. Of further interest was evidence showing significant variation in risk levels across contexts for high defensive boys--more cautious under "adult evaluation" and more risky under "peer competition"--and the absence of such variation for low defensive boys. (Project conducted in collaboration with Mr. J. Damm of the University of Queensland, Australia.)

Cognitive Styles and Attitudinal Consistency (NICHHD, 784-9) Mr. Messick

PURPOSE: Some of the major concerns of social psychology in recent years have dealt with the dynamics of attitudinal consistency--in particular, (a) with the extent to which both cognitive and affective components of attitude systems tend to be organized in psychologically consistent patterns, (b) with the operation of motivational pressures toward consistency, and (c) with the specific mechanisms whereby inconsistency is resolved. The present study investigates the importance of reliable individual differences in the tendency toward attitudinal consistency, the generality of their operation across different measurement approaches and different response processes (such as interpersonal perception and opinion change), and the extent of their relation to other cognitive and personality variables.

PROGRESS: Attitudinal consistency was measured using four experimental procedures that had previously formed the basis for four major studies in this area: The McGuire Logical Consistency Test, the Morrissette Balance Completion Test, the Osgood and Tannenbaum Congruity Test, and the Rosenberg and Abelson Fenwick Test.

These four experimental procedures were administered, along with a measure of intolerance of attitudinal inconsistency, to a sample of 200 high school seniors who had been extensively tested earlier with a battery of cognitive and personality measures. One ancillary function of the present study will be an evaluation of the replicability with high school students of the mean effects observed in the four previous experiments by Morrissette, McGuire, Rosenberg and Abelson, and Tannenbaum. The major emphasis, however, is upon the consistency of interrelations among these four procedures, upon the number and nature of dimensions required to account for individual differences in attitudinal consistency, and upon the cognitive and personality correlates of those dimensions. Particular attention will be paid to the relations between categorizing styles and attitudinal consistency, since level of conceptual differentiation and cognitive complexity may moderate consistency pressures. (Study done in collaboration with J. E. Singer of State University of New York at Stony Brook.)

Social Stratification (NICHHD, 784-16) Mr. Stricker

PURPOSE: A vast amount of research has been devoted to the identification and development of objective indexes of social stratification, such as education and occupation. Although this research has been based almost entirely on samples of whites, these indexes have often been used to measure the stratification of Negroes, despite suggestions that these indexes are not comparable for the two groups.

An inherent problem with these indexes, even when they are restricted to whites, concerns the identification of the cutting points for differentiating between the social classes. Empirical attempts to identify these points have largely been unsuccessful and a theoretical controversy continues over whether the social classes are distinct entities, reflecting discontinuities in stratification, or are arbitrary classifications that have been imposed on a continuum.

This study has two aims: (a) to identify the dimensions of social stratification for both Negroes and whites, and the indexes that best reflect these dimensions, and (b) to determine whether distinct social classes exist and, if so, whether they can be empirically differentiated.

PROGRESS: A comprehensive, highly structured interview schedule was prepared, focused on information that directly reflects the major aspects of social stratification or are important correlates of it. The schedule was used in household interviews with representative samples of Negro and white household heads in a typical northern city. The data are being analyzed.

Training of Cognitive Functions in Children (NICHHD, 786-1)

Mr. Jacobs

PURPOSE: A procedure is being developed to train the inductive reasoning ability of children. It will be used to investigate the increments in ability possible at different age levels and different initial ability levels.

PROGRESS AND MAJOR FINDINGS: The double-classification problem forms the basis of many intelligence test items, and is of major interest in Piagetian theory. The solution of such a problem requires *S* to take into account simultaneously two different aspects of stimulus variation (e.g., size and shape) while inferring a logical relation. Two previous studies attempted to increase the scores of first-grade children on Raven's *Coloured Progressive Matrices*, a test that includes many double-classification items. The rigorously prescribed training procedures used resulted in a temporary item-specific effect (1966 *Proceedings of the 74th Annual Convention of the American Psychological Association*, pp. 263-264), and a heightened response set (*Perceptual and Motor Skills*, 1968, *27*, 759-766). The insights gained in these studies suggested that more general and lasting improvement in the skill of handling double-classification problems could be obtained by altering certain features both of the training and of the criterion.

As a by-product of the two above studies, it was found that a score based on type of distractor chosen on the CPM test has a moderate degree of test-retest reliability, concurrent and predictive validity, and cross-cultural applicability (*Psychological Reports*, 1970, *26*, 311-315).

A new training procedure was developed that allowed for more individualization of instruction and involved more verbalization. *Coloured Progressive Matrices* was replaced by a new test of double-classification skill which (a) used shapes which children could label easily as the elements of each matrix, (b) was constructed in three randomly parallel alternate forms, and (c) permitted the measurement of transfer along specifiable dimensions. Experimental *Ss* scored significantly higher than control *Ss* on both learning and transfer tests, both immediately and when tested for retention four months later with M. Vandeventer, *Child Development*, 1971, *42*, 149-159.

The general methodological questions involved in relating training operations to operations used to measure transfer are discussed in "Evaluating the training of intelligence." This paper maps out a universe of relations within which transfer could be assessed more meaningfully (RB-69-20, with M. Vandeventer).

A subsequent research study (1) assessed transfer within this universe, (2) evaluated the effects of more extensive training, and (3) investigated the effectiveness of different trainers. As in the earlier study, trained *Ss* significantly outperformed control *Ss* on shape and color matrices. Their superior performance also transferred throughout the universe of relations. More extensive training produced significantly more transfer than regular training. There was no difference in the effectiveness of the two trainers (RB-69-88, with M. Vandeventer).

A study recently completed and now being written up shows that under some circumstances one can obtain transfer across *operations* as defined in Guildford's Structure-of-Intellect model.

Learning to Formulate Original Responses in Relation to Anxiety and Defensiveness (NICHHD, 786-10)

Mr. Frederiksen
Mr. Evans

PURPOSE: This study involves an attempt to improve performance on the Formulating Hypotheses Test (which is intended to measure "creativity" by requiring subjects to develop hypotheses to account for data) by a kind of programmed instruction involving feedback.

PROGRESS AND MAJOR FINDINGS: Results of a preliminary study showed that: (1) the training increased quantity but not the quality of ideas produced; (2) the improvement did not transfer to a similar task requiring divergent production; (3) test anxiety was curvilinearly related to performance, with poorest performance associated with an intermediate level of test anxiety; and (4) there was a significant interaction of anxiety and verbal ability. Results were interpreted in terms of the effects of training on setting standards of performance rather than altering ability (Journal of Educational Psychology, 1969, 60, 465-475).

The study now in progress involves two kinds of feedback materials, one emphasizing quantity and the other emphasizing quality of ideas. The purpose of this investigation is to see if quality as well as quantity of ideas can be improved and to see if the relationships of anxiety and verbal ability to the production of ideas are influenced by the nature of the feedback materials. Sex differences will also be investigated. Data have been obtained from about 400 college freshmen from two eastern state colleges. New methods for scoring protocols to yield scores reflecting quality as well as quantity have been developed. All scoring has been completed, and analysis of data is in progress.

Are Bayesian Optimal Decisions Artificially Intelligent? (NICHHD, 786-12)

Mr. Alker
Mrs. Hermann

PURPOSE: Claims have been made that humans, in comparison with Bayesian formulas, are essentially conservative in extracting the maximum permissible certainty from probabilistic information. Two alternative hypotheses concerning observed discrepancies between human subjects and Bayesian formulas are considered: (1) the Bayesian model simulates well the information processing of intelligent, abstract, rational, nondogmatic subjects. Unintelligent, concrete, irrational, dogmatic subjects, with demonstrated difficulties in processing probabilistic information, on the other hand, will exhibit the expected reluctance to revise their opinions in the light of new information, (2) applying the Bayesian formulas to complex, important human decisions oversimplifies those problems thus accounting for humans' apparent conservatism.

PROGRESS: An initial report of the findings was presented at the 1967 APA meetings. The final report is available (RB-70-27) and is in press in the Journal of Personality and Social Psychology.

MAJOR FINDINGS: (1) No individual difference variable of previously demonstrated relevance to accuracy in processing probabilistic information related to conservative deviation from optimal Bayesian value. A measure of risk-taking preference also was unrelated to this purported human conservatism. (2) A manipulation of problem type, which subjects reported successfully, increased the importance and complexity of the problem, increased subjects' deviation from optimal Bayesian performance in a conservative direction. The negative findings from (1) and positive findings from (2) suggest not that humans are inherently conservative but rather that Bayesian formulas, as simulations of human decision making, are inherently risky.

Factors in Comprehension of Verbal Material
(NICHHD, 786-14)

Mr. Carroll
Mr. Freedle

PURPOSE: From its inception, this project has been envisaged as a program of separate studies to determine various parameters in the comprehension of discourse and in the learning of facts, concepts, and principles from such discourse. Particular attention would be paid to the role of lexicon and syntactical structure. Although some of the work would be done with native speakers of English responding to oral and written discourse in English, studies would also be concerned with the comprehension of discourse in the learning of a second language in order to examine the development of comprehension in detail. Also, comprehension ability would be studied developmentally by conducting experiments and tests with individuals sampled from different age levels, particularly the years of schooling.

PROGRESS: During 1970-71, this project contributed toward the completion of the monograph, "Learning from Verbal Discourse in Educational Media," that was submitted to the U. S. Office of Education in January 1971.

Mr. Carroll organized a Research Workshop on Language Comprehension and the Acquisition of Knowledge held March 31 - April 3, 1971 at the Quail Roost Conference Center, Durham, N. C.; Mr. Freedle assisted him as administrative coordinator of the conference. The conference was supported, in the main, by the Committee on Basic Research in Education of the National Research Council in cooperation with the National Academy of Education. Papers were presented, with subsequent discussion, by 16 well-known research workers from the United States, Canada, and England. Messrs. Carroll and Freedle gave two of the papers. This conference is regarded as closely related to the purposes of the project being conducted under the Center grant. It is planned that the proceedings of the conference will eventually be published as a book, to be edited by Messrs. Carroll and Freedle.

Simulation of Learning Processes in a Restricted Set of Experiments from the Literature (NICHHD, 786-15) Mr. Carroll

PURPOSE: It is becoming clear that the number of variables that may be involved in any given learning situation and the number of basic learning paradigms that may apply are so large that any complete analysis of learning situations could only be done by complex computer programs. The goal of this research is to fashion a taxonomy of learning processes such that the outcomes of learning situations can be more adequately predicted and controlled, and to test this taxonomy by applying it to clusters of learning experiments reported in the literature.

PROGRESS: Work on this study has been held up during the past year because of the director's activities on other projects.

Application of the Lognormal Model of Word Frequency Distribution (NICHHD, 786-16)

Mr. Carroll

PURPOSE: In previous work, it was demonstrated that the lognormal distribution yields good fit to word frequency distributions when account is taken of the bias in drawing finite samples from a theoretical population. The lognormal theory makes possible the estimation of the type-token function and the vocabulary size underlying a given sample. The research aims to explore the application of the lognormal model to various types of samples in order to aid in the characterization of individual differences in verbal ability and the description of samples of verbal material.

PROGRESS: Work has continued on the application of the lognormal theory to further data that have become available from various sources, and on the preparation of a monograph to report the results.

The theoretical orientation of this project has led to explorations of several topics connected with word frequency, its measurement, and its correlates:

(1) A new and apparently more satisfactory formula for assessing the dispersion of words over samples has been devised, using information-theoretic statistics. Furthermore, the principal investigator has proposed a new scale for expressing word-frequencies, somewhat on the analogy of the familiar T-scale. This scale, called SFI (for "Standard Frequency Index"), computes SFI by the formula $SFI = 10 (\log_{10} p + 10)$, where p is the probability (proportional frequency) of a word in a corpus. For example, words with a frequency of one per million are assigned SFI = 40. This work is reported in RB-70-48, and a paper based thereon has been accepted for publication in Computer Studies in the Humanities and Verbal Behavior, entitled, "An alternative to Juilland's U, and a proposal for a Standard Frequency Index (SFI)."

(2) A further investigation of the measurement properties of subjective magnitude estimates of word-frequencies was made (following B. J. Shapiro, "The subjective estimation of relative word frequency," Journal of Verbal Learning and Verbal Behavior, 1969, 8, 248-251). Fifteen lexicographers and 13 other adults gave subjective magnitude estimates for 60 words. The reliability of the averaged z-score ratings was .9900 for the former, .9744, which are significantly different values. For the two samples, the respective correlations with objective frequency data were .970 and .923, again significantly different. When objective frequency is held constant, the partial r between the two sets of estimates was .657, a value suggesting that subjective estimates are independently valid measures of word frequency that may be more accurate than objective data, which are biased through inadequate sampling. A manuscript reporting these results is nearly ready for publication.

(3) As noted in last year's report, the rationale that has been developed for the asymptotic lognormal distribution has certain resemblances to a theory of retrieving words from memory that was developed by R. C. Oldfield. To explore this problem, data were collected on naming latencies for 94

pictures of objects from 50 subjects. Analysis of the data is still under way, but it is clear now that word frequency is not as relevant to the naming latencies as another type of variable, namely, the "age" of the word in memory. Regardless of frequency, objects that have names that were acquired very early in life (_cup_, _giraffe_) are named much faster than objects that have names that were acquired somewhat later in life (_tent_, _metronome_). This result suggests that long-term memories are laid down and retrieved in accordance with the age at which they were acquired. A manuscript reporting these results is in preparation.

Studies of the Acquisition of Grammar in a Foreign Language (NICHHD, 786-17)

Mr. Carroll

PURPOSE: The basic problem is to analyze how the foreign language learner acquires competence in the grammatical structure of the language, and to resolve possible conflicts between two types of theory concerning this process--the "audiolingual habit" theory and the "cognitive code-learning" theory.

PROGRESS: During the period covered by this report, progress has been held up because of activities on other projects.

Generalizability of Cognitive Styles (786-18) Mr. Cahen

PURPOSE: The purpose of the research is to study different areas of cognitive style and to develop theoretical models and techniques for assessing group as well as individual behavior in these areas.

PROGRESS: Data are being analyzed from an administration of Kagan's Matching Familiar Figures Test. The data were obtained under two conditions. The first condition obtained measures where the students were asked to select a distractor which was the same as the presented stimulus. The second administration was obtained under conditions where the students were asked to pick a distractor that was different from the stimulus. Measures of response latency and errors under the two conditions are being examined and compared. A preliminary analysis of the data indicates that the dimension of latency is more stable within and across measures than the dimension of errors. Additional MFF data from other sources and investigators are being gathered and analyzed to check the stability of the measures of latency and errors.

Additional sets of MFF data have been obtained from three investigators in the United States and Canada. These data are being analyzed for the relative contribution of latency and error to the construct of impulsivity.

Human Choice Behavior (NICHHD, 786-19) Mr. Freedle

PURPOSE: The goal of this project is to establish an axiomatic basis for the analysis of behavior generated in a variety of experimental settings such as paired-associate learning, human probability learning, syntactic and phonemic recognition studies, and more general psychophysical sensory discrimination studies. Establishing the lawful relations among stimulus presentations and their associated responses will serve as a basic step in designing and analyzing more complex studies in human learning and cognition.

PROGRESS: Three approaches to the problem of choice behavior are now in progress.

(1) The analysis of synthetic asymmetric confusion matrices using a modified input to the Shepard-Kruskal computer scaling program indicates that this program will recover the specially constructed "double" configurations which are asserted to account in part for asymmetric matrices. Application of this modified program to phonemic recognition data suggests that the distinctive features of "nasality" and "voicing" are either partially attenuated or are totally lost (i.e., not sampled) on a proportion of the experimental trials thereby contributing to the observed asymmetry of the data matrices. An analysis of asymmetry in syntactic recognition is available as RB-70-40. The study on phonemic recognition is still in progress.

(2) General Choice Theory

a. Luce's choice axiom, when it is recast in terms of proportion of time spent in making, say, manipulative or observing responses, appears to provide one way to study its applicability to an infant's attentional behavior and also provides a means of indirectly determining whether the infant perceives subgroupings of specially constructed multidimensional stimuli. This study is available in Developmental Psychology, 1970, 2, 129-133. A second theoretical paper is available in Developmental Psychology, 1971, 4, 240-247.

Further quantitative approaches to infant behaviors involving vocalization interaction with his mother is available in Merrill-Palmer Quarterly (in press). Also a quantitative analysis of play behavior from the point of view of unfolding analysis was presented at a Symposium on Genetic and Social Influences on Psychological Development, July 4-8, 1971 at University of Nijmegen, The Netherlands.

b. A second aspect of general choice theory concerns a modification of Luce's theory in a different direction from that above. This revision deals with a quantitative account of the effects of "irrelevant" stimuli and anchor stimuli on response probabilities for choices made from the available set of responses. A paper on this topic was given at the Psychometric Society Meetings held at Stanford University, March 21-22, 1970.

(3) Human Probability Learning. Two studies testing an assumption concerning the subject's memory of the last k events and the effects this information exerts on his subsequent choices were conducted. A paper dealing with this result is available as RB-69-4. The results of the second study were presented at the Mathematical Psychology Meetings, Stanford University, August 28-29, 1968.

MAJOR FINDINGS: The analysis of several asymmetric matrices has suggested that the probability of creating dimensional information which is not present in a stimulus is significantly smaller than the probability of losing dimensional information which is present in a stimulus. Also the probability of not sampling stimulus cues dealing with syntactic structure leads to asymmetric matrices.

The evaluation of the memory assumption for k events (an assumption made in a learning model known as the k-span learning model) indicates that a subject performs best on patterns of k events which involve few runs and performs worst on patterns involving many runs (with the exception of those patterns which involve a regular alternation of events). This indicates that modifications of the k-span model which reflect the results of the above studies are called for.

The more general choice theory dealing with the effects of "irrelevant" stimuli on choice probabilities indicates that a better fit can be obtained within the current theory as compared with Luce's choice axiom; however, in addition, it is shown that the mathematical structure of Luce's choice equations are obtainable as special cases of the general choice model.

The quantitative investigations which have been applied to developmental problems suggests that a number of infant behaviors (such as toy preference attention to visual stimuli, response decrement or "practice" curves produced by repeated presentation of visual stimuli, etc.) can be integrated by Luce's concept of response strength measures, where this measure has now been defined to apply to proportion of manipulation time, or proportion of viewing time instead of the relative frequency approach used in Luce's earlier work.

Attentional Distribution as a Function of Novelty and Incongruity (NICHHD, 786-21)

Mr. Lewis
Mrs. Wilson
Miss Van Wyk
Mr. Weisman

PURPOSE: We are interested in determining if there are attentional differences to stimuli that are defined as novel as opposed to those defined as incongruous. By incongruous we mean an alteration in relatively familiar material such as a cat's head on a horse's body. By novel we mean material not experienced before. This definition of novelty is difficult to interpret in terms of real events in that a stimulus is usually multidimensional and some aspects are certain to be familiar. In the present experiment we are calling a picture of an anteater novel. From one theoretical position we could predict that familiar stimuli (cat or horse) would elicit less attention than an incongruous stimuli (cat's head-horse's body or horse's head-cat's body) and that a novel stimulus (anteater) would elicit the most. Alternative theories would argue that novel and familiar stimuli would be more equal in attention-eliciting behavior and would be attended to less than the incongruous stimuli. Because a priori classification of stimuli as novel, familiar and incongruous is difficult at best, the children's verbal labeling of the stimuli will be used. Thus, if one child responds with the word bear to the picture of the anteater, and another has no label we predict that these different verbal responses will be associated with different attentional behaviors.

PROGRESS: Fifty 3 1/2- to 5 1/2-year-old subjects have been run and the data have been analyzed.

MAJOR FINDINGS: The results suggest that attention, verbal labeling and latency to respond with a label are highly related. A Research Bulletin is available (RB-70-56).

Learning of an Alternating Stimulus Pattern in Young Children (NICHHD, 786-21)

Mr. Baskett
Mr. Lewis

PURPOSE: The purpose of this investigation was to study the orienting response to a familiar stimulus. Prior research has shown that if young children are presented a single visual stimulus over a series of trials, eye fixation, vocalization, smiling and other responses to the stimulus change in a systematic manner, generally showing a decrement in attention. However, if on a later trial a new stimulus is presented, the subject's attention will be reinstated. For this study, the two stimuli were presented in an alternating pattern (e.g., A-B-A-B-A-B) and after a series of repetitions of the stimulus pair, the sequence was broken or violated (e.g., A-B-A-B-A-B-B). This experiment was designed to ask whether or not the subject learns the stimulus in the context of prior stimuli. If the subject had learned the pattern, he would have been expected to show greater attention to the violation since he had never experienced B following B. The response measures were duration of the first eye fixation, total fixation time, smiling and vocalization as recorded by an observer and body movement which was recorded mechanically. The stimuli were two figures, one of 20 straight black lines and the other of 20 curved colored lines and both were recorded on slides.

PROGRESS: Twenty-one children, between 3 1/2 and 6 years old, were assigned to three groups of seven subjects each and individually exposed to the repeating pairs of stimuli. The intertrial interval was .5 seconds. The groups experienced either 2, 4, or 8 pairs of stimuli before the violation occurred, and then, with the violation being the first element of a new pair, either 7, 3, or 1 repetitions of the pattern respectively.

MAJOR FINDINGS: Since the data for the visual fixation was clearest, and since both smiling and vocalization were very infrequent, only the fixation data will be discussed. As a result of a large within subject variability, no significant differences were observed across trials for the three groups. Comparisons were made across the groups for the trials on which each group received its first or second violation (for all three groups) and the third and fourth violations (for two groups). The differences, tending to be significant, generally reflect the fact that for a given violation the groups had experienced a different number of trials. That is, attention was greater for those groups having experienced fewer trials. Keeping the number of trials constant, the differences on the violation trials seemed to be due to the group which had received eight violations. This group tended to be more attentive, perhaps because they received a more varied stimulus pattern. Also, this group tended to look at the violations longer than the nonviolation trials. The other two groups did not exhibit this phenomenon. One final finding was that the subjects tended to prefer the colored stimulus more than the black and white stimulus. In general, the results suggest that if the hypothesis is true about the violation of the pattern, then the effect builds up very quickly, in this case after only two repetitions of the pattern.

Longitudinal Study of Infant Attention and Its Relationship to Subsequent Perceptual-Cognitive Development
(NICHHD, 786-21)

Mr. Lewis
Mrs. Wilson

PURPOSE: This longitudinal study will investigate attentional processes during infancy and their relationship to cognitive growth in the preschool period. Specifically, the study will explore: (1) problems in measuring attention; (2) factors influencing attention, such as sex, state (awake or asleep), and clinical problems (birth trauma); (3) stimulus differences such as facial vs. nonfacial and familiar vs. novel; (4) developmental consequences of differential attention, including the relationship of attentional differences in infancy to preschool performance in concept formation tasks, learning problems, reading, and language acquisition.

PROGRESS: To explore these problems, two samples of 32 infants, 16 boys and 16 girls, were seen at 6, 13, 25 and 44 months of age. Sixteen of the subjects were also seen at 9 and 18 months. At each age level, a variety of visual and auditory stimuli were presented to the infants and a large number of response measures were taken. In addition to their response to visual and auditory stimulation, a wide variety of other behaviors were observed. These included such behaviors as concept formation, problem solving, task persistence, language acquisition, play behavior, and IQ measures.

The data, yet to be fully analyzed, will be used to explore the major areas of interest outlined below.

(a) Measurement of attention. The responses observed have been: (1) fixation (receptor orientation), (2) vocalizing, (3) crying, (4) autonomic reactivity (specifically, heart rate and respiration) and (5) gross motor activity. Investigators often harbor implicit assumptions about each of these response indices. For example, differential fixation time is often assumed to reflect a capacity for differentiation or discrimination. More significantly, absence of differences in fixation time is believed to reflect the infant's inability to differentiate among the visual patterns being presented. This inference may be fallacious, for often the infant fixates for an equal period upon two stimuli but will only vocalize or smile to one of them. Vocalizing has been assumed to indicate both a specific social response and general arousal level. Similarly, gross motor activity or heart rate acceleration is often regarded as an indicant of activity or arousal level. It is too early to comment on the validity of these assumptions, but it is clear that detailed empirical work is necessary to assess their viability. At present, it appears that each of the infant's responses viewed singly is probably ambiguous in meaning and can reflect different psychological processes. A long fixation time can occur to a familiar stimulus or to an uncertain one. Cardiac deceleration can occur when the infant studies the stimulus and is quiet, but acceleration often appears when he studies the stimulus and is active. Thus the present study intends to contribute information relevant to understanding the varied responses associated with distribution of attention to auditory or visual patterns. Analysis of six-month data indicates that attention to visual and auditory stimuli is usually associated with the following type of response pattern: receptor orientation; decrease in gross motor activity, crying, and vocalizing; autonomic responses of cardiac deceleration; and a slowing of respiration. Individual differences in this response pattern have been observed. Developmental differences in this response pattern have yet to be explored.

(b) Individual differences in attention. Marked individual differences have been found in the attentional behavior of six-month-old infants. Three types of individual differences will be explored: (1) attention span (total time attending); (2) stimulus differentiation (discrimination); and (3) habituation. Individual differences both to visual and auditory stimuli will be observed over the four age levels. Moreover, these differences will be related to other variables, including the following: (a) early mother-infant interaction as measured by an observer both in the home and in the laboratory; (b) socioeconomic level of the parents; and (c) physiological factors such as birth condition as measured by Apgar Score and birth records.

(c) Stimulus differences in attention. Because a variety of visual and auditory stimuli were presented at each of the four age levels, it will be possible to explore the influences of different stimuli on attentional behaviors. Both visual and auditory stimuli can be divided into two classes: (1) social and nonsocial, i.e., human forms, faces, or voices versus nonhuman patterns or sounds; and (2) novel and familiar, i.e., distortions of human form, syntax, and vocabulary versus their regular nondistorted forms. These two classes of stimuli were presented at all age levels and the study seeks to explore the effects of these stimulus qualities on such attentional processes as their ability to attract (length of time oriented toward) and hold (amount of habituation) the infant's attention.

(d) Relationship between attentional measures and other cognitive capacities and styles. Along with the various attentional measures to visual and auditory stimuli, a variety of other cognitive capacities and styles were observed for the one-, two-, and three-year levels. These include: (1) language tests at two and three years, (2) play behavior at one, two and three years, (3) concept formation at two and three years, (4) problem solving at two and three years, and (5) IQ scores at three years. The cognitive style measures of impulsivity and reflectivity as measured by MFF task and persistence task were given at three years. It will therefore be possible at each age level to investigate the relations among these various measures of cognitive capacity and to relate them to the attentional measures. For example, there is evidence that amount of fixation is related to play behavior at one year and that IQ scores can be predicted from attentional scores at the three-year level.

(e) Predicting cognitive capacity and style in the three-year-old from early attentional differences. One of the major aims of the study is to relate early attentional behavior to subsequent cognitive growth. It was hypothesized that early advantages in attentional ability should facilitate subsequent development. For example, infant girls appear to show significantly better auditory discrimination than infant boys, which might account for the more rapid language development in girls at later ages. It is clear that the ability to discriminate sounds successfully must precede the production of sound. Individual differences in preschool behavior will be related to early attentional differences. Preliminary results on one attentional variable indicate that the Stanford-Binet IQ score at three years of age can be predicted from the amount of habituation to visual stimulation at one year.

(f) Personality development. In the longitudinal investigation of attention and cognitive growth, several personality characteristics were assessed. Observation of individual differences in attentional behavior revealed that personality characteristics might covary with them. Personality measures obtained at the one-, two- and three-year levels were: (1) measurement of the infant's vigor, (2) dependence behavior as measured by interaction of infant with mother or teacher, (3) sex role appropriate behavior measured in play behavior, and (4) task persistence. Preliminary observations suggest that a dependent child may spend less time looking at nonsocial stimuli as compared to social stimuli. Moreover, when left alone in a relatively unfamiliar room, his attention span might be more a function of wanting to see the mother than of interest in the stimulus being presented. It was also observed that vigorous children had more difficulty attending than placid, less vigorous ones who showed little physical activity. Analysis of the data is proceeding along all domains of interest.

The Mother-Infant Interaction in Africa (NICHHD, 786-21) Mr. Lewis
Miss Lusk

PURPOSE: In our studies of mother-infant interaction we were led to believe that one of the most crucial aspects of that relationship is the contingency relationship between the infant's action and the mother's response. We have hypothesized that this contingency interaction is crucial in developing a motive of efficacy such that the infant feels that his action is capable of causing consequences.

In order to explore this we have started some pilot cross-cultural work in Senegal, Zambia, Holland and the United States.

PROGRESS: Data have been collected in Zambia and Senegal and are being analyzed.

MAJOR FINDINGS: We have observed mother-infant contingency interactions and measured infant cognitive capacity with the hope of demonstrating that the stronger the contingency relationship, the faster the cognitive capacity--cognitive capacity being measured in terms of attention differences and in terms of some infant mental scales such as the Bayley.

On Relating an Infant's Observation Time of Visual Stimuli with Choice Theory Analysis
(NICHHD, 786-21)

Mr. Freedle
Mr. Lewis

PURPOSE: The purpose of this note is to outline a method by which observation time as measured by visual fixation on an object in the visual field can be predicted for n -items viewed simultaneously in the field, when the time for observing each item presented singly is known. Problems of item categorization are considered for their potential importance in inferring the perceptual-cognitive process by which an infant selects stimuli to attend to. Further theoretical work indicates that a stimulus similarity scale can be constructed using attentional measures. Also mathematical models of learning and preference behavior (as in toy-preference behavior) are under way and to date appear to give a very good fit to empirical data.

PROGRESS: A theoretical article dealing with the choice among n-alternatives is available in Developmental Psychology, 1970, 2, 129-133. A second theoretical paper on a stimulus similarity scale is available as RB-69-89 and in Developmental Psychology, 1971, 4, 240-247. Research reports on learning applications and toy-preference behaviors which use a Mathematical Modeling Approach are currently in progress.

Social Class and Sex Differences in the Attachment and Play Behavior of the Year-Old Infant (NICHHD, 786-21) Mr. Lewis

PURPOSE: In this experiment lower class white infants, approximately 15 girls and 15 boys, were seen in a free play situation exactly like that of Goldberg and Lewis.

PROGRESS: The report was issued as a Research Bulletin 70-13. It has been accepted for presentation at the Eastern Psychological Association meetings and has been accepted for publication by the *Merrill-Palmer Quarterly*.

MAJOR FINDINGS: Response to their mothers and their play behaviors were observed. The most outstanding social class difference was that both lower class boys and girls showed significantly less vocalization to their mothers than did the middle class infants. This result suggests that the linguistic deficit in the lower class environment can be seen in the year-old infant.

Sex differences in attachment behavior were also observed and the results replicated the Goldberg and Lewis study. Girls showed greater attachment to their mothers than did boys and the results indicate that differential socialization in appropriate sex role behavior is consistent across class.

What's Thrown Out with the Bath Water: A Baby? (NICHHD, 786-21)

Mr. Lewis
Ms. Johnson

PURPOSE: To investigate possible sample bias.

MAJOR FINDINGS: Although the exact figure is not usually reported, it is estimated that up to 30 per cent of the infants who participate in experiments within the first half-year of life fail to provide valid data and must be discarded as subjects. There are a variety of causes for this; most of them might be classified as state problems such as sleepiness, fatigue, or fussiness. The exclusion of such large numbers of subjects would, in animal or questionnaire research, be looked upon as having a potentially biasing effect on the reported data. No such assumption is made in human infant research. Indeed, it is implicitly stated that infants excluded are exactly like those used except for the fact that at that particular time of testing they were "uncooperative."

Whether this assumption is correct required confirmation, for it may be possible that these excluded infants constitute a special class of subjects and the inclusion of their data would greatly alter the results. The present investigation was a first attempt to explore this sampling problem by observing the data available from "incomplete" subjects and comparing it to "complete" subjects seen at the same time.

Data from infants unable to complete the experimental sessions were compared to those for whom there were complete data. The results suggest that the elimination of large numbers of infants may have a potentially biasing effect on reported data. A report of the study was issued as Research Bulletin 70-35 and has since been accepted for publication in Child Development.

Mrs. Darley
Mr. Lewis

Activity and the Cardiac Response (NICHHD, 786-22)

PURPOSE: Some theorists believe that the cardiac response is directly related to the subject's intended transaction with his environment, such that deceleration occurs when the subject wants to "take in" information and acceleration occurs when the subject wants to exclude stimulation.

Alternatively other theorists believe that the cardiac response is one of several response systems which are inhibited when the subject attends to his environment. The theories differ in that the former holds the cardiac response control to the attentional response while the latter holds that it is only peripheral to other activity-inhibiting responses.

PROGRESS: In order to investigate this problem we are looking at data already collected and investigating fixation, activity and cardiac responses. We hope to show that partialling activity out of fixation trials results in greater deceleration than nonfixation trials. Moreover, partialling fixation out, deceleration will be greater for nonmovement than for movement trials.

The study is a first attempt at getting a fit on this important theoretical problem. The data are still being analyzed.

A Developmental Study of the Cardiac Response to an Auditory Signal: Onset and Offset of a Stimulus
(NICHHD, 786-22)

Mr. Lewis
Mrs. Ban
Mrs. Baumel

PURPOSE: The present study is concerned with the cardiac response of infants within the first year to the onset and offset of a complex musical stimulus. What different information processing is involved when a stimulus goes on than when it goes off and how should this affect the cardiac response? Moreover, what is the effect of repeated presentation of the same signal? Finally, what age differences in information processing to an auditory signal exist over the first year?

PROGRESS: The data indicate several important results: (1) a consistent and rather large degree of deceleration; this for both onset and offset. (2) For 3- and 6-month-olds the terminal orienting response (OR) is less than the onset OR, while the 12-month-olds show greater terminal than onset OR. (3) Habituation of the onset OR follows a developmental pattern, with 3-month-olds showing the least habituation and 12-month-olds showing the most. (4) Habituation of terminal OR also shows a developmental trend, with the 3- and 6-month-olds showing habituation while 12-month-olds showed little habituation. The recovery data results for the seventh trial indicate that some changes are more effective than others. A paper was presented at the tenth annual meeting of the Society for Psychophysiological Research in New Orleans in November 1970. A Research Bulletin is available (RB-71-5), and is in press in Psychophysiology.

Developmental Studies of Children's Incidental Learning and Distractibility (NICHHD, 786-23) Mr. Hale

PURPOSE: This research looks at children's performance on various measures of attention. These include assessments of (a) distractibility, (b) incidental learning, and (c) dimension preference. The research examines the validity of theories which postulate developmental changes in selective attention, as measured by these paradigms. Interrelations among these measures are also assessed.

One of the major purposes of the research is the development of a measure of distractibility. The paradigm involves a short-term memory task, present under different distraction and control conditions. Various types of visual and auditory distractors are included, and differences in performance on distraction and control conditions reflect the influence of each type of distractor.

PROGRESS: Data collection has been completed, and the results of the portion of the research involving distractibility are reported in a manuscript, entitled "Five- and eight-year-old children's response to auditory and visual distraction" by the project director and Edward Stevenson, which has been submitted for publication as a Research Bulletin. Preliminary analyses of the remaining data have been performed, and these results will be reported in connection with subsequent research.

MAJOR FINDINGS: Both auditory and visual distraction were found to impair performance of children of both ages 5 and 8. The degree of impairment was roughly equal for both age levels, although a marginally significant effect was observed for visual distraction; the detrimental effect of visual stimulation tended to be greater for the 5-year-olds than the 8-year-olds. The main results concerning incidental learning were that (1) the 8-year-olds showed a greater amount of incidental learning than the kindergarten children and (2) overtraining had essentially no effect on the amount of incidental learning shown at either age level.

Human Choice Behavior in Psycholinguistic Studies Mr. Freedle
(NICHHD, 786-25)

PURPOSE: The purpose of this project is to examine the lawful (presumably probabilistic) relationships between the usage of a certain class of words, such as prepositions, nouns, or adjectives and the physical characteristics and situational settings of objects presented in a controlled experimental setting. Three modes of attack are currently under way all of which indicate that an adequate choice model to describe the complexity of the behavioral data must be formulated at the level of probabilistic (stochastic) models. The first study relates preposition usage in response to a variable object A which assumed a variety of positions around some fixed object B; a second study has dealt with the types of messages which subjects write in order to communicate information about a target item with respect to four other similar items in an array. This second study indicates that a probabilistic model of noun and adjective usage is definitely appropriate and furthermore provides evidence for an extension of a cognitive theory of reference due to David Olson. The third approach has constructed an information-processing model of syntactic stimuli; again, a stochastic model was seen to be necessary to provide a good fit to the data.

PROGRESS: The results of this last study were presented at the Mathematical Psychology Meetings, Miami Beach, Florida, Sept. 1-3, 1970. The analysis of the preposition usage study is still in progress. A report on the analysis of messages (with Philip Kingsley) is in preparation.

A paper which developed a number of quantitative models of the comprehension process was presented at a Research Workshop on "Language Comprehension and the Acquisition of Knowledge" held at Rougemont, North Carolina, March 1971. Included in this paper was a discussion of the probabilistic approach to theories of reference as well as latency models (information-processing models) of the comprehension process.

MAJOR FINDINGS: The application of an information-processing model to the data collected by what is known as a sentence "probe" technique has indicated that at least three stages can be identified in determining the time it takes subjects to search through short-term memory so as to give a correct response with respect to information in some stimulus sentence. The first stage was identified as the time it takes to locate in short-term memory the "probe" word position, the second stage was interpreted as the time it takes to perform the task of searching around the "probe" word in memory so as to retrieve the correct information, and the third stage was identified as a fixed time constant, the time it takes to ready the vocal apparatus and then give the response. Previous models of "probe" data have failed to give a complete account of the distribution of latencies; the present stochastic model gives a very good account of the entire set of data.

The probabilistic approach to a cognitive theory of reference indicates that appropriate models can be found which handle both individual differences within experimental treatments as well as differences across treatments; special cases of this general probabilistic approach to reference yield Olson's algebraic model of reference.

Stimulus Clues in the Judgment of Veracity
(NICHHD, 786-25)

Mr. Freedle
Mr. Baskett
Mr. Kingsley

PURPOSE: People are often placed in the situation of having to make a judgment as to whether or not another person is telling the truth. Many examples are available to indicate the generality of this situation, such as when a parent interrogates a child or when lovers question each other, or in the courtroom when testimony is being presented and challenged. The purpose of this investigation is to examine certain clues that people might use to arrive at these judgments. Subjects were asked to listen to a tape recording of an adjective checklist which is being responded to by another person who has been instructed to answer falsely on some of the adjectives. Two such tapes were used, with each adjective being answered affirmatively and negatively.

PROGRESS: Two tapes were prepared with nine latency levels on each tape. The subject was asked to decide whether or not the stimulus person was lying or not and how certain the subject was of his judgment. Two experiments were conducted to examine what factors influenced those judgments. The first experiment involved 83 paid subjects who listened to the tape recordings. In the second experiment, the stimulus adjectives were rescaled as to the degree to which people (in this case, 71 college students) felt the adjective was true of other people.

MAJOR FINDINGS: Overall, it was found that if the target person responded either too quickly or too slowly, the subjects attributed his response as a lie more often than if the delay was more intermediate in duration ($p < .01$). The adjective likability value in combination with the response of "true" or "false" by the target person also contributed to the attribution of lying ($p < .01$). These results are in preparation. The second study indicated that the likability of an adjective and the believability of the adjective are not highly correlated and that the degree of believability is related to the degree to which lie attributions are made. Stemming from this research, a research model for studying the perceptions of lying was proposed (RM-70-12).

Difficulty Factors in Learning to Spell Mr. Cahen
(NICHHD, 786-26)

PURPOSE: The purpose of this research is to identify stimulus properties of spelling words that are associated with the spelling difficulty of the words.

PROGRESS: The literature on spelling difficulties has been reviewed and published as Research Bulletin 69-72. A modified version of this report is in press in the Review of Educational Research. A data bank which will be utilized to identify stimulus properties of spelling words that are associated with spelling difficulty of the words has been built. Data (proportion of students passing a given spelling word) from the New Iowa Spelling Scale will form the dependent variables in the analyses. A set of linguistic variables has been constructed for the predictor matrix. A Research Bulletin (70-47) has been issued. This report provides a mathematical model for the prediction of spelling difficulty. A draft of a second Research Bulletin has been prepared, and this reports a more sophisticated regression model using interaction terms. A shortened version of this paper was presented at the 1971 meeting of the American Educational Research Association. A first set of tables of phoneme-grapheme and grapheme-phoneme correspondences has been developed. The orthographic system of Professor Richard Venezky (University of Wisconsin) has been used and the frequencies have been weighted by the data from the Brown University Study (Kučera-Francis, principal investigators). This first set of tables has been weighted by types. It is planned that the correspondences will also be weighted by tokens. A comparison of the prediction utility of these two different types of tables will be compared. Both sets of correspondences have been broken down for the initial, medial, and final position in a word.

In a second study, an experiment was run which will add information about the effect of the semantic usage of words on spelling difficulty. The experiment utilized fourth- and fifth-grade students. Two forms of the spelling test were created and administered to matched groups of students. One form contained a word in one context with the same word in another context being assigned to the second form. For example, the word listed was given in one spelling test in the following format: "The boy listed the words on his paper." The second form utilized the following sentence: "The boat listed from side to side." The spelling tests were given orally and the data are now being transcribed from tape to coding sheets and eventually will be on punched cards for analyses. Differences in proportion of students passing the two words in the different contexts will be studied to determine the relative contribution of spelling ability, the general frequency of the word in the language, and parameters of context as determined from the Lorge semantic count list on spelling performance.

Interrelations in Children's Learning of Verbal and Pictorial Paired Associates (NICHHD, 787-27) Mr. Hale

PURPOSE: It has been found that children learn paired-associate tasks more easily when the tasks employ pictorial rather than verbal materials, and when the materials are concrete rather than abstract. The present study sought to determine the degree of relationship among these various tasks in the processes by which they are learned. Children in grades 3, 6 and 9 were given a number of paired-associate learning tasks, which involved materials which were either verbal or pictorial, concrete or abstract. Intercorrelations among learning and retention scores on these tasks were determined.

PROGRESS: The project has been completed and the results are reported in Research Bulletin 71-4.

MAJOR FINDINGS: A task involving pictorial materials was significantly easier to learn than a task involving the verbal representation of those pictures for children in grades 3 and 9 but not in grade 6. The task involving concrete words was easier to learn than a task involving abstract words for children in grades 3 and 6 but not in grade 9. The correlational analysis indicated that the task involving pictures and the task involving the concrete words representing those pictures were more highly correlated at grades 6 and 9 than at grade 3; for all other pairs of tasks, the correlations remained relatively constant across grade levels. This result is consistent with evidence from an earlier study and is interpreted to suggest that the processes by which children learn the picture and concrete word tasks become more similar as the child grows older.

Manipulated Expectancies and Teacher Bias Mr. Cahen
(NICHHD, 787-28)

PURPOSE: The purpose of the study is to examine teacher bias and manipulated expectancies and correlates of these behaviors.

PROGRESS: Data are being analyzed to understand further the psychological processes associated with teacher bias or influenceability as conceptualized in the study. Preliminary analyses indicate that the amount of bias contributed to manipulated expectancies has been found to be a function of the amount of information given to a teacher to establish the expectancy. A draft of a Research Bulletin dealing with this analysis has been written. A second analysis will be performed on the data that have been gathered. The second analysis will focus on the personality dimensions of teachers which are associated with the amount of bias displayed in the scoring of pupils' papers. The second analysis will be sequential in nature. It will include separate analyses at different stages of scoring a set of four tests. Normative data for the tests used in the study have been obtained under controlled conditions. These data were obtained on approximately 250 subjects.

Matrix Sampling and Abilities (NICHHD, 787-29) Mr. Cahen

PURPOSE: The purpose of this research is to study empirical conditions affecting item or matrix sampling. The research will also focus on the mathematical properties of the technique in determining the relationship of the precision of measures to the number of students tested under the technique.

PROGRESS AND MAJOR FINDINGS: A report (RB-68-39) was published in the Spring 1970 issue of Educational and Psychological Measurement. This study reported the results of a large-scale item-sampling administration with ninth-grade students. An additional set of item-sampling data for twelfth-grade students has been analyzed and is reported in a Research Bulletin (70-64). This Bulletin has been submitted for publication. The analyses from the second study have given further indication that item-sampling is an efficient method for estimating school means and that the amount of error in the estimation is a function of the number of students tested in the sampling unit.

Developmental Changes in Self/World Differentiation (NICHHD, 787-30)

Mr. Baker

According to certain leading developmental theorists (e.g., Piaget, Werner, Schachtel), boundaries between self and world are lacking in human infancy and evolve only with higher levels of development. In particular, it is assumed that the relationship between self and world can vary from one of nondifferentiation (i.e., fusion or linkage) which prevails in infancy to one of relative differentiation (i.e., separateness, distance from, or oppositeness) which characterizes the adult. Several interrelated studies have attempted to demonstrate that even among normal adults, degree of self/world differentiation can vary, and that such variation has measurable effects upon perception. It is now proposed to extend this work (1) by studying age changes in perceptual phenomena, using children as subjects; and (2) by studying individual differences in perceptual experience among adult subjects along one or more dimensions of self/world differentiation.

Study 1

PURPOSE: In studying age changes in perceptual phenomena, the specific dependent variable chosen for study is the psychophysical error of anticipation (= undershooting) which has long been known to occur in space localization experiments when some variable stimulus, presented in an ascending or descending trial, is moved toward some standard value. The central hypothesis which shaped our earlier studies with adults--viz., that with increase in degree of self/world differentiation there is an increase in magnitude of undershooting--was derived in part from developmental theory. As such, it would seem a most natural extension of this work to explore directly whether there are changes during ontogenesis in the effects of these variables which the earlier research demonstrated affects adult performance. It is thus proposed to undertake a series of studies, cross-sectional in design, using children between the ages of 7 and 19.

PROGRESS: Moderate progress has been achieved to date in the following areas: (1) construction of a laboratory for conducting the planned perceptual-cognitive experiments was completed in January 1971; (2) necessary equipment has been constructed and is now set up in the laboratory; (3) an extensive series of studies has been sketched out; (4) some pertinent literature review has been accomplished; (5) the running of subjects in the initial research phase has commenced; (6) the more elaborate and refined equipment which will be necessary in the second phase of this research has been designed; (7) a well qualified full-time research assistant, with some expertise in the area of perception, was hired and trained in a variety of dark-room laboratory procedures which are necessary for conducting the planned research; (8) data from some published studies by other authors have been procured and analyses have been undertaken to clarify issues relevant to this project; and (9) reports of four studies completed by the present researcher prior to the initiation of this project, but pertinent to its broader purposes, have been prepared; one report has been issued (RB-70-45).

Study 2

PURPOSE: The present study has as its focus individual differences in the degree of self/world differentiation among normal adult subjects. Although a number of studies in the past decade have demonstrated that variation in the relationship between self and stimulus object does have measurable effects upon perceptual experience, for the most part such studies have been experimental in nature--that is, conditions were varied, and the goal was to demonstrate that some general effect resulted from this variation. Studies of individual differences in the ongoing relationship between self and object have been quite rare and almost always limited to a single perceptual situation. The present research seeks to go beyond this by attempting--for the first time--to explore whether there are consistent individual differences in self/world orientation which can be assessed by a battery of tests tapping a variety of perceptual and cognitive situations.

PROGRESS: Work on this project, which was initiated prior to the researcher's joining the ETS staff, was resumed during the summer of 1970. Progress has been achieved to date in the following areas: (1) space has been procured and a laboratory built at the Psychology Department of Florida State University for conducting the planned perceptual-cognitive experiments; (2) all equipment has been designed and is now in operational use; (3) plans have been finalized regarding the initial phases of this project; (4) an extensive set of materials were prepared specifying in detail the procedures, instructions, ways of recording data, etc. for each of the tasks decided upon; (5) the person who will be collecting the data during the initial phases was trained in the appropriate laboratory procedures; (6) to insure that standard dark-room lab procedures would obtain, the present researcher spent two days at the Florida State University laboratory checking out all of the procedures, and (7) an initial study involving part of the total proposed battery has been conducted. This involved the standard rod-and-frame test; a similar test in which a rod alone is presented to the subject; two variants of the rod-and-frame test--one with instructions which make a person highly aware of his own body during the task, and another in which the person is asked to suppress all information from his body while carrying out the task (these latter two based on some research by Baker, 1968); and a novel task, using the rod-and-frame apparatus, in which the subject is asked to line up the stimulus rod with the longitudinal axis of his tilted body (after Baker, 1963). It is hoped to launch research with the full battery during the summer of 1971. (Study done in collaboration with Seymour Wapner from Clark University and Joseph Grosslight and Deena Hullman of Florida State University.)

Newspaper Tests (ANPAF, 788-1) Mr. Diederich

PURPOSE: The American Newspaper Publishers Association Foundation sponsors a "Newspaper in the Classroom" program that reaches over 3 million students in grades 7-12 with 68 million copies of newspapers per year. The Foundation sought ETS help in developing tests to measure some of the effects of this program.

PROGRESS: Eight tryout forms of tests based on two simulated newspapers of four pages each, tabloid size, were administered in a spiral order to 4,000 students in 61 schools. Half the students were in newspaper classes, the other half in control classes of equal reading ability without newspaper training. Items on which newspaper classes surpassed control classes were selected for pre- and posttests for junior and senior high schools, published by Cooperative Tests and Services.

MAJOR FINDINGS: Average differences between newspaper and control classes on the 30 items selected for each test were significant at the .01 level. The types of competence revealed by these items were classified in 30 categories that correspond to most of the objectives of newspaper programs and define them more clearly than ever before.

The project is now in the final stage of establishing norms for newspaper classes and for classes without any special training in reading newspapers in grades 7-12.

Infant Development in Lower-Class American Families (Childhood Council--NICHHD, 789-1)

Mr. Lewis
Mrs. Sarett

PURPOSE: This study is directed toward observing differences in mother-infant interaction as a function of socioeconomic differences.

MAJOR FINDINGS: Over 35 children, three months of age, from all five SES categories as designated by Hollingshead were subjects in the study. Two hours of observation of the interaction was recorded in the infant's home under as natural a setting as possible. A wide variety of behaviors were observed, employing the same procedures used in other studies. The data have been collected and reveal both differences in amount of stimulation and nature of stimulation, for example, lower class mothers in general provide more stimulation to their infants than middle class mothers. Moreover, there are significant differences in responsivity as a function of SES. When an infant vocalizes, a middle class mother is more likely to vocalize back than a lower class mother. These differences were discussed in terms of social class differences at later ages. A paper reporting on this study was presented at the Society for Research in Child Development meetings held in Minneapolis in April 1971.

Mother-Infant Interaction and Infant Development among the Wolof of Senegal (Childhood Council, 789-1)

Mr. Lewis
Miss Lusk

PURPOSE: Caretaker-infant interaction within the first year of life was studied in a group of 10 Wolof infants.

MAJOR FINDINGS: Using observational techniques as well as standard testing procedures, several results emerge. The pattern of caretaker-infant interaction was more strongly related to age of infant than any other variable investigated. As infants become older their behavior changes from predominantly passive-reflexive to active-operant, while their caretaker's behavior switches from proximal to distal. The often-found result that African infants show precocious development within the first year was confirmed for the Wolof. No relationship between the interaction measures and tests of cognitive and motor skills was found. A report has been prepared and issued as Research Bulletin 71-14, and is in press in Human Development.

Mothers and Fathers, Girls and Boys: Attachment Behavior in the One-Year-Old (Childhood Council, 789-1)

Mrs. Ban
Mr. Lewis

PURPOSE: Recent studies of attachment behavior have focused primarily upon mother and child interaction. A variety of measures has been employed in an attempt to isolate specific "modes" of attachment behavior; when possible, these modes have been discussed in terms of sex and age differences and social class influences. The bulk of these studies has revealed that attachment behavior cannot be generalized for all children, but assumes many forms which are comprehensible when the above-listed variables, as well as others, are considered. The present study attempts to utilize the question of sex differences in regard to attachment, along with that of another variable which has so far received little attention--attachment to father. Thus, in a design to be outlined, boys and girls were compared in their modes of attachment behavior to their fathers as well as to their mothers.

MAJOR FINDINGS: That sex differences do exist in regard to attachment to mother has been reported in previous studies. However, not so clearly defined is the behavior directed to the father who, though normally a central figure in the child's environment, nevertheless plays a vitally different parental role from that of the mother. It is the dual purpose of this study to further the knowledge of attachment behavior and, through this, to better define some of the differences in the relationship of the mother and the father to the child. Ten boys and ten girls, along with their mothers and fathers, were used as experimental subjects. The age of the children was one year, ± two weeks. The socioeconomic background of the sample was predominantly upper-middle and professional class.

The room in which the subjects were observed has been described in detail in an earlier paper. All subjects paid two visits to the observation room. One visit was made with the mother and the other with the father. In order to control for any visit effect, both boys and girls were divided into two groups, such that half made the first laboratory visit with the mother, and the other half first with the father. Each observation period lasted for 15 minutes. During this time, the parent sat in a chair in a corner of the room while the child was free to roam about at will. An observer who stood behind a two-way mirror depressed a key on an event recorder each time one of the following behaviors occurred: the child was in close proximity to the parent, he touched the parent, he vocalized to or looked at the parent. In addition, the latencies of the latter three measures were noted.

Means and standard deviations were computed for the seven measures. These suggest that not only do significant sex differences exist in the expression of attachment generally, but also that the mode of attachment behavior is influenced by the sex of the parent to whom it is directed. Both boys and girls spent significantly more time in close proximity to, and touching, their mothers than their fathers. Similar, but not quite as strong, differences were found for amount of vocalizing to the parent.

Analyses of variance further confirm the differences outlined above and reveal information about the strength of contribution of the variables introduced: sex of the child, sex of the parent, and visit number. An analysis of variance on amount of touching parent reveals that the greatest amount of touching is directed to the mother on the second visit, then to the mother on the first visit, then to the father on the second visit, and, finally, to the father on the first visit. An analysis of variance on amount of vocalizing reveals a nearly identical pattern. Other analyses reveal similar differences as well as interaction, including visit, sex, and parent effects. In summary, the results confirm that boys and girls express attachment behavior differently to their mothers and that this behavior, in turn, is different from that directed to their fathers.

The results of the analyses outlined above uniformly affirm the questions which guided the study. They are: Are there sex differences in attachment behavior which can be revealed through the seven measures employed? Is there a difference in the mode of attachment behavior directed to the father when compared with that directed to the mother? These affirmations require further investigation into the reasons for such differences. That is, how much of the sex difference in attachment is due to genetic control, and how much to differential treatment by the mother? Further, what are the implications of the differences in attachment to the parents? Clearly, the behavioral differences suggest that the child perceives his relationship to his father as different from that to his mother. But does the greater intensity of attachment to the mother, insofar as touching, proximity, and vocalizing are concerned, simply reflect the greater amount of time normally spent in association with the mother, or are there other variables, including genetic, operating? As is so often the case with behavioral studies, answers simply evolve into further questions. Whatever the case, the present study suggests new dimensions in the area of attachment, and introduces further variables which appear to influence attachment behavior. A report on this study was presented at the Eastern Psychological Association meeting in New York in April 1971.

Stability of Attachment Behavior: A Transformational Analysis (Childhood Council--NICHHD, 789-1)

Mr. Lewis
Mrs. Ban

PURPOSE: This study constitutes a systematic investigation into three major aspects of a child's attachment to mother: its mode of expression (i.e., the interrelationships of those behaviors which can be defined as attachment-motivated), sex differences in regard to these specific behaviors, and, finally, the consistency of expression of these behaviors across age.

MAJOR FINDINGS: A sample of 30 males and 30 females was observed at 13, 25, and 44 months in a series of controlled settings. The settings used to observe attachment fall under one of the following three descriptive categories: free play, observed at 13, 25, and 44 months; separation from mother, observed at 13 and 25 months; and the presence of a stranger, observed at 25 months only. The data collected from these observations, as well as such demographic information as socioeconomic class, enabled us to undertake statistical analyses which would highlight the nature of attachment as a function of such variables as age, sex, socioeconomic class, and experimental condition. A brief summary of the results of the completed analyses follows.

Both sex and age differences in regard to attachment behavior during the free-play episode have been observed. When amount of time in proximity to mother, amount of touching mother, amount of time looking at and vocalizing to mother are compared, significant sex differences are seen within age and significant age shifts are in evidence. At 13 months, girls exceed boys in the amount of expression of all four attachment behaviors. At 25 months, the sex differences are generally reduced, except for vocalizing, although both sexes vocalize significantly less to mother at 25 months than at 13 months. Age shifts for two of the behaviors produce sex by age interactions. Females touch mother significantly more often than do boys at 13 months, but a decrease in the amount of female touching behavior at 25 months eliminates the sex difference at that age. A similar age change in amount of proximity also occurred.

Of primary importance in the analysis of the attachment behaviors discussed above is the nature of the interrelationships noted among them. That is, for both sexes at 13 months, the touching and proximity behaviors correlate highly, while vocalizing and looking also correlate highly. Further, these two groups of measures appear to have a generally negative relationship and have been termed proximal and distal attachment behavior respectively. This behavior dichotomy does not hold as strongly at 25 months, for the correlation between the two groups of measures is positive for males and inconclusive for females. Thus, the proximal-distal attachment behavior dichotomy tends to fall away by 25 months. Most notable in the across-age relationships of these measures is the fact that a significant across-age "reversal" occurs in regard to the proximal attachment measures for males. That is, the proximally attached boys at 13 months are those who are least attached to mother at 25 months. All four of the attachment behaviors at 25 months are generally negatively related to the 13-month proximal attachment behaviors for males. Such dramatic behavior shifts reveal further the age-specific nature of developmental trends.

Separation from mother was accomplished by the use of a barrier situation. Analyses of the children's behavior reveal a trichotomy which was observed at both the 13-month and 25-month age levels for both sexes. A component analysis was completed for a total of 10 barrier separation behaviors. At both ages, the analysis suggested three general modes of response: on the one hand, a subgroup of children responded through active barrier approach, with specific attempts to surmount the barrier and with little or no sign of overt distress. A second component suggested a subgroup who expressed heightened distress through uncontrolled crying, with little active attempt to surmount the barrier. A third component suggested a subgroup characterized by passivity who neither approached the barrier nor showed signs of distress.

At 13 months, for boys, the barrier separation trichotomy relates systematically to the proximal-distal dichotomy observed in attachment behavior. In a component analysis incorporating the 10 barrier separation and four attachment measures, the component distribution was as follows: boys expressing heightened distress at separation were those boys who were proximally attached, while boys who showed little or no distress and made active and systematic attempts to surmount the barrier were those who were distally attached.

The relationship of the attachment-in-the-presence-of-a-stranger section to the first two sections remains to be analyzed. Generally, the analysis thus far completed suggests strong sex differences in regard to attachment behavior, as well as significant age changes, even within the first few years of life. A paper reporting on this study was presented at the Society for Research in Child Development meetings held in Minneapolis in April 1971. A larger monograph is in preparation.

Factorization of Covariance Matrices in Terms of a Specified Pattern or Structure (NSF, 791-1)

Mr. Kristof
Mr. Wingersky

PURPOSE: Let Σ be a given covariance or correlation matrix of order $n \times n$ and P a specified matrix of order $n \times m$ and full column rank, $m < n$. The basic problem solved here is the determination of a symmetric nonsingular matrix $\hat{\phi}$ of order $m \times m$ and a diagonal matrix $\hat{U}$ of order $n \times n$ such that, in a least-squares sense, $\Sigma = P \hat{\phi} P' + \hat{U}$. Hence we are seeking a set of uniquenesses in $\hat{U}$ and a factor variance-covariance matrix $\hat{\phi}$ that are maximally compatible with Σ and P.

PROGRESS: The work was presented at the Psychometric Society Meeting, 1971. A Research Bulletin is in preparation.

MAJOR FINDINGS: Solutions for $\hat{\phi}$ and $\hat{U}$ are obtained in closed form without the use of differentiation. A computer program has been written. The method has been applied to a problem of factor comparison in a cross-cultural setting.

Generalization of the Orthogonal Procrustes Rotation Procedure to More Than Two Matrices (NSF, 791-1)

Mr. Kristof
Mr. Wingersky

PURPOSE: When more than two factor matrices are to be conjointly compared, paired comparisons proliferate and complicate interpretation. To control this tendency a method is derived, given a number of factor matric[illegible]or determining one "average" factor matrix and a set of orthogonal mat[illegible]s rotating the "average" factor matrix toward the given factor matric[illegible]

PROGRESS: The work will be presented at the 79th Annual Convention, APA, 1971. A Research Bulletin is in preparation.

MAJOR FINDINGS: The "average" factor matrix on rotation by the orthogonal matrices will approximate the given factor matrices in a least-squares sense over the full set of given factor matrices. The procedure has been programmed for digital computer in FORTRAN and applied to date.

Some General Results on Fit in Factor Rotation Mr. Kristof
(NSF, 791-1)

PURPOSE: A rather general theory of oblique factor rotation is outlined.

PROGRESS: The work is contained in RB-70-60 and has been accepted for publication in Psychometrika.

MAJOR FINDINGS: The main results are formulated as four theorems. Necessary and sufficient conditions are derived for two factor matrices to admit identical factor structures and/or factor patterns with factors having unit variances. These conditions are expressed in terms of eigenvectors and eigenvalues of certain matrices obtainable from the data. It is also shown that two matrices admitting identical factor structures will admit identical factor patterns and vice versa. After introducing the notion of a pair of transformations to identical structures and/or identical patterns, rules are given as to finding such pairs if they exist. Finally, some immediate consequences of the theorems are noted. They concern, for example, the suitable choice of a target structure and/or pattern and a hierarchical order of jointly necessary and sufficient conditions for fitting a specified target perfectly.

Evanston Integration (Rockefeller Foundation, 795-2) Mrs. Hsia

PURPOSE: Elementary schools of Evanston, Illinois were desegregated by redistributing and bussing in 1967. A longitudinal assessment of the impact of desegregation upon pupils, teachers, parents, and the community was undertaken jointly by District 65 and ETS.

PROGRESS: Data have been collected and analyzed. A final report will be completed in July 1971.

MAJOR FINDINGS: Desegregation has had little effect upon the academic and personal growth of black and white children. Parents, teachers, and the community remain committed to a principle of integrated schooling. There has been considerable increase in community interest in participation in school related activities.

International Study of Educational Attainment in French as a Foreign Language (USOE through Teachers College, Columbia University, 798-19) Mr. Carroll

PURPOSE: This is a part of the International Study of Educational Attainment (I.E.A.) being conducted by an International Council headed by Professor Torsten Husén of the University of Stockholm, and in which achievement tests in English as a Foreign Language, French as a Foreign Language, Civics, Literature, Reading Comprehension, and Science will be administered in a number of countries for the purpose of comparative studies. The French as a Foreign Language tests are being monitored by an international committee of which Mr. Carroll is chairman and Mr. Clark is a member. Tests of listening, speaking, reading, and writing are to be prepared for various levels of achievement, ranging from that of a 10-year-old population to that of the pre-university level.

PROGRESS: The tests approved by the International Council of the IEA Study in December 1969, with minor modifications, were administered in a number of countries in the spring of 1971 to large, statistically controlled samples of students in grades 10 and 12. A set of tests were also administered to 10-year-old pupils in the United States. The administration of the tests was handled by national centers in the various countries.

In the course of devising procedures for analyzing data, Dr. Carroll made a study of methods of scoring speaking and writing fluency tests. A Research Bulletin, RB-70-52, pointed out that this type of test may require the use of nonlinear functions such as ratios or proportions in order to control for the amount of productivity.

Final analysis of the data will be performed mostly in Stockholm, Sweden, in the year 1971-72.

U. S. National Testing in French as a Foreign Language (USOE through Teachers College, Columbia University, 798-20) Mr. Carroll

PURPOSE: This project comprises all work connected with Project 798, Job 19 that concerns the pretesting and eventual final administration of tests for French as a Foreign Language in the United States. The work is monitored by a National Committee on French as a Foreign Language.

PROGRESS: No work was done, or required to be done, under this project during 1970-71 since all responsibility for the national testing in the United States rested with the national Center at Teachers College.

Recall of Approximations to English (799-2) Mr. Fremer

PURPOSE: The relationship between hearing-deafness, age, mastery of formal language, and recall was studied using various approximations to English (AEs), that is, lists of words whose structure and content may be viewed as falling between that of randomly ordered word lists and connected prose.

PROGRESS: Multiple-choice recognition tests were prepared for measuring amount of recall of AEs that had been previously developed as part of this project. The recognition tests were pretested in January and February 1969. Final tests were administered in May and June 1969 to hearing children in grades 4, 7, and 11 and to deaf children at three age levels spanning the elementary and secondary school years. A report of the study was issued in March 1971 (TDR-71-1).

MAJOR FINDINGS: All six of the groups tested, three deaf and three hearing, showed some evidence of greater recall of structured or higher order approximation to English lists over their recall of unstructured or first order approximation to English lists. The hearing children showed increased recall from first to second, second to third, and third and fifth to seventh order of approximation. The two oldest deaf groups showed improvement in performance from first to second and from second to third order of approximation, but showed no improvement beyond that point. The youngest deaf group showed improvement in performance from first to second order of approximation only. Because of this failure to show improvement in performance on lists with greater contextual constraint, all three groups of deaf children performed substantially less well than hearing children on higher orders of approximation, even though the two older deaf groups performed as well as or better than the hearing groups on the first order lists.

The findings on number of words recalled suggest that the deaf children do not possess a fundamental grasp of the connectedness of language that is commensurate with their ability to recall unrelated words and their knowledge of vocabulary. Although they can recall the constituent words, they demonstrate only limited ability to profit from the language structure that permits hearing subjects to remember words in groups or clusters.

A Study of the Interrelationships among Reading Comprehension Ability, Writing Ability, and the Linguistic Structure of the Written Language of Sixth-Grade Children (799-9)

Mrs. Massad

PURPOSE: The purpose of this investigation is to determine if significant interrelationships can be shown to exist between the written language ability and the reading comprehension ability of sixth-grade pupils. In so doing selected receptive and expressive language functions are being examined in detail through a reading comprehension test, an objective test for writing ability and an analysis of the linguistic structure of children's written language as expressed in free essay. As a pupil's ability to read and write is an indication of the most noticeable effects of language instruction provided by the schools, an understanding of the interrelationships between the abilities to read and write could provide educators with guidelines for instruction.

PROGRESS: One hundred eighty sixth-grade children were tested. The STEP Reading and Writing Tests and a free essay were the data-gathering instruments used. All computerized data analyses have been completed, and a report on this phase of the study is nearing completion. A more extensive analysis of the linguistic structure of the children's language as expressed in the free essay is still in progress. Results of this more extensive analysis were not included in the computerized data analyses due to a very high level of interdependence among the factors being studied.

MAJOR FINDINGS: A factor analysis of the variables indicated that the instruments involved measure (a) reading comprehension ability, (b) ability to use conventions of language--such as rules of grammar and organization patterns, (c) ability to select appropriate language, and (d) two aspects of the linguistic structure of freely written language--the number of communication units and the number of words per communication unit. Of particular interest is the finding that these two aspects of the children's freely written language are not significantly related to the children's reading comprehension ability nor to the children's ability to use the conventions of language to express themselves. (A portion of this study is being done in collaboration with Mr. Kender of Lehigh University.)

Study of Academic Performance
(Western College Association, 801-9)

Mr. Warren

PURPOSE: The increasing practice of using college degrees as credentials on which the award of occupational and social positions are based makes more critical an understanding of the kinds of capabilities that college degrees reflect. The student behaviors associated with faculty judgments of high academic achievement will be described with attention to their diversity and consistency across fields of study, academic levels, and individual faculty members.

PROGRESS: About 200 faculty members at 12 California four-year colleges have described the ways a good student and a poor student in a particular class differ from a mediocre student in the same class. Content analyses of these descriptions are being conducted.

Cambodia Aftermath Survey (Carnegie Commission on Higher Education, 809-1) Mr. Peterson

PURPOSE: The purpose of the survey was to describe the national campus reaction to the May 1970 Cambodia incursion and the Kent and Jackson State killings.

PROGRESS: A questionnaire was sent to all the college presidents in August of 1970. Seventy-three per cent responded. All analyses and report writing have been completed. A book-length report is scheduled for publication in September of 1971 by the Carnegie Commission. A mimeographed report has been completed (R. E. Peterson, Cambodia, Kent, Jackson and the Campus Aftermath, October 1970).

Academic Credit by Examination (CASE, 839-1) Mr. Sharon

PURPOSE: The focus of this study is on the nontraditional student who has obtained part of his education outside of a formal school system and who was admitted to a college on the basis of the Tests of General Educational Development (GED). In addition to determining the validity of the GED test battery for predicting the college success of the nontraditional student, the problems and satisfactions of these candidates will be assessed. A variety of assessment techniques, including tests, questionnaires, and interviews are being employed.

PROGRESS: Twenty-nine colleges and universities have provided data on 1371 students who have been admitted with GED scores. Detailed questionnaires have been completed by many of these students. Interviews with a small number of students were conducted on the campuses of four institutions.

System of Interactive Guidance and Information (Carnegie, 842-2) Mr. Katz

PURPOSE: A computer-assisted system of interactive guidance and information is being developed to improve career decision-making by students in junior colleges.

SIGI is a guidance system which allows the student to interact with a computer in such a way as to examine his own values, obtain and use relevant information, interpret predictive data, and formulate plans. This interaction assists the student in arriving at tentative career decisions and in modifying them as he gains new insights and additional information. The decisions involve both educational and occupational options. Emphasis, however, is not merely on the content of decisions but on the process of decision-making. As the student progresses through SIGI, he learns to move freely within the structure of the system. In gaining control of the system, he progressively develops competencies and masters strategies for rational behavior in the face of uncertainty.

PROGRESS: Content has been outlined, scripts have been written and programmed, a system has been designed, "Mark I" student terminal has been built, a computer language has been developed, demonstrations have been run, and students have interacted with the system. The script-writing, terminal development, and programming have now gone through an iterative procedure, with each component being successively improved, refined, and articulated with the others.

Assessment and Development of Cognitive Skills in Young Children: A Classroom Approach (Carnegie, 858-6)

Mr. Chittenden
Mrs. Bussis
Mrs. Orost
Mrs. Tanaka

PURPOSE: Initiated in July 1970, the central purpose of this two-year project is to complete development of materials that will help teachers understand and foster the intellectual development of children in prekindergarten through grade three. Goals are to provide the teacher with specific guides and techniques to help observe, interpret, and encourage significant behaviors in children between the ages of approximately 4-9 years old. The final materials will incorporate and thus supercede the current Let's Look at First Graders material.

PROGRESS: Tasks and materials for the early grade levels are nearing a final stage of completion and are based largely on previous research phases of this project which were also supported by Carnegie. During the past year, additional measures have also been developed and tried out at the 1st, 2nd, and 3rd grade levels. Major emphasis is being placed on four aspects of development: language functioning and communication skills; quantitative thinking; classification and ordering; use of egocentric and intuitive processes. In addition, attempts will be made to appraise certain style variables as well as other behaviors (self-confidence, self-awareness) more commonly classified under the heading of "self-concept."

Development of Procedures for Assessing Open Education in Follow Through Programs (USOE, 858-50)

Mr. Chittenden
Mrs. Amarel
Mrs. Bussis
Mrs. Tanaka

PURPOSE: The first purpose is to state the basic objectives of open education (as represented by the Education Development Center Follow Through Program) in a form which provides a framework for evaluation. The second purpose is to begin the development of assessment procedures which are appropriate to such an educational position.

PROGRESS: During the first year of the project (July 1969 to June 1970), observations were made of EDC Follow Through classrooms in several different states; interviews with EDC advisory staff were completed; and some new measures were given preliminary trials with kindergarten and first-grade subjects. During the second phase of the project (July to December 1970), preliminary developmental work proceeded on measures to appraise certain cognitive and affective characteristics of children at the third-grade level. Areas selected for test development were: intuitive judgments; quantitative thinking; self-perception; communication and group problem solving; and language abilities.

MAJOR FINDINGS: A report of the first year's work was completed (PR-70-13). This report contains a conceptual analysis of the EDC approach to open education, as well as implications for research and evaluation. A detailed report of the second phase of the project, together with recommendations for further development of selected measures, was submitted to the United States Office of Education in December 1970. No further work for the Follow Through Office of OE has been undertaken since December.

Study of Teacher Change in Open Educational Settings
(865-19)

Mr. Chittenden
Mrs. Bussis
Mrs. Amarel

PURPOSE: This project is a teacher interview study addressed to what are considered two major facets of educational experimentation and accountability: (1) effective implementation of an educational approach, and (2) the creation of an evironment that will sustain and promote continuing growth within that approach. Initially, at least, our concern with these questions will be focused on "open" educational environments and research on the teacher's role--her educational priorities, perceptions, attitudes, working assumptions. Specific goals of the project are as follows: (1) to identify and describe concrete issues and factors which teachers perceive as both facilitating and inhibiting of change toward an open classroom; (2) to link characteristics of teachers, their environment, and their perceived needs for assistance into a better theoretical understanding of the dynamics of change; and (3) in the course of the research, to refine an interview instrument and set of scoring procedures which would be applicable to teachers working in innovative reform programs other than open education.

PROGRESS: Work on this project began in late April. To date, a preliminary interview has been drafted and administered to 15 teachers. During the summer, it is anticipated that more interviews will be administered and an initial working version of the interview and scoring procedures will be constructed.

A Study of Intellectual Growth and Vocational Development (USOE, 870, All Jobs)

Mr. Hilton
Mrs. Casserly
Mr. Evans
Mr. Freeberg
Mr. Godwin
Mr. Katz
Mr. Patton
Mr. Rock
Mr. Rosenfeld
Mr. Warren

PURPOSE: The objectives of the study are: (1) to trace the intellectual development of students who subsequently elect vocational as opposed to college preparatory curricula; (2) to investigate the interaction over time of stated vocational plans, individual characteristics, environmental influences, and subsequent occupation; (3) to develop a preliminary theoretical model of vocational development.

PROGRESS: The work is being conducted as a series of separate but related studies. Job 10, the Vocational Decision-Making Study, involves the construction and pretesting of a structured interview schedule to collect data on vocational decision-making. The data include individuals' perceptions of occupational values, strategies for seeking and processing information, and rationale for present occupational position or plans. A description of the interview schedule is contained in RM-69-3.

Job 11, the Curriculum Group Achievement Study, involves a multivariate analysis of variance of the antecedents and achievement of students enrolled in various high school curricula. The dimensions of the design are school, curriculum, sex, and socioeconomic status. RB-69-29 reports the results of this study.

In Job 12, Simulation of Curriculum Assignment Process, steps are being taken towards a computer simulation of the process by means of which high school students become associated with a particular academic curriculum.

In Job 13, the Follow-Up Study, follow-up questionnaires were sent to graduates of Growth Study schools. Analysis included grouping the subjects by sex, high school curriculum, and current activity; observing the relationship between curriculum and current activity; and then comparing the pattern of activities since high school and plans for the future across the various groups. For one or two Growth Study school systems post-high school activities in relation to race and perhaps social class will also be observed.

The purpose of Job 14, the School Setting Study, is to assess by means of interviews the "climate" of the Growth Study schools and the communities in which they are located in order to gain data to explain observed differences in test performance, dropouts, and curriculum enrollments in the various Growth Study schools. Staff members visited most of the Growth Study schools to conduct interviews with counselors, teachers, and students, using an interview schedule developed for this purpose.

Job 15, the Biographical Factor Stability Study, involves the analysis of the responses to the Background and Experience Questionnaire, which was administered to the total Growth Study sample in 1963, 1965, and 1967, to determine the factor stability over time of the responses and changes in validity of the factor scores based upon STEP scores as criteria.

For Job 16, Moderated Prediction within Curriculum, approximately 900 students from one Growth Study school system having complete information on aptitude tests, background questionnaire, curriculum grouping and criterion information were selected for extensive analysis in an attempt to define and describe those measurable characteristics of an individual which may interact with his achievement and his curriculum. RB-69-32 describes the results of this study.

The Dropout Study, Job 17, is an investigation of the pattern of early academic achievement, background, and academic growth as it relates to school persistence. The first phase of this study focused on one large Growth Study school system which has participated in the Growth Study since 1961. In the second phase the study was repeated in a second school system. A Research Bulletin is in preparation.

The purposes of Job 18, Negro-White Differences, were (1) to investigate the intellectual growth patterns of Negro and white students as reflected by their scores on SCAT and STEP at four different points in time (grades 5, 7, 9, and 11) and (2) to utilize the Background and Experience Questionnaire and information on curriculum enrolled in to describe the cultural patterns associated with the differences. Racial data were collected from the two largest Growth Study school systems and were put on the data tape. RB-69-71 describes the results.

MAJOR FINDINGS: Cutting across all the studies are the following generalizations:

1. Each of the multivariate analyses of variance and the on-site school studies point to one major generalization, mainly, that one cannot generalize across schools. Although there are general trends and overall differences, each school is unique--in the structure and organization of its programs, in the performance of its students within each curriculum, in the performance of students grouped by race, or by sex within each curriculum.

2. In all the studies it has been necessary to recognize the strong and pervasive relationship of each student's family status to his choice of curriculum and to his performance within that curriculum. To understand school effects, curriculum effects, and racial effects, it is essential to understand the precise nature of the impact of family status.

3. Large scale multivariate analytical techniques, although powerful, must be supplemented with other data collection and analytical methods. These techniques serve largely as methods of data reduction and to identify significant differences and interactions. Explaining the differences and interactions requires further analysis and, often, additional data collection, sometimes from field studies and observation at a clinical level. Copies of the 409-page final report may be obtained from Mr. Hilton, R-244.

Test and Job Performance of Various Subgroups within Specified Occupations (Ford Foundation, 879, Jobs 2 & 3)

Mr. Campbell
Mrs. Crooks
Mr. Rock
Mrs. Mahoney
Mrs. Parry

PURPOSE: This project seeks to investigate the relationship between job and test performance for different ethnic groups when pertinent background factors are taken into consideration.

The study grew out of a concern by the U. S. Civil Service Commission regarding the impact of tests on the selection and advancement of minority group members. The objective of this study is to identify tests which are good predictors of job performance and at the same time do not place certain cultural subgroups at an unfair disadvantage. The project is being carried out under a grant from the Ford Foundation as a joint responsibility of ETS and the U. S. Civil Service Commission.

PROGRESS: The first job studied, medical technician, was reported in the 1968-69 Summary. The 1969-70 Summary contained reports of instrument development and test administration for the second job studied, cartographic technician. Data gathered from the cartographic technicians was analyzed during 1970-71. Work was started in the summer of 1970 on the third job studied, inventory management specialist, and visits to a number of government installations were made to study job duties. Twelve aptitude tests were selected which appeared to measure abilities related to the job, a personal history questionnaire was constructed, and a set of supervisors' rating scales on 10 job dimensions was designed. A work sample, in the form of an in-basket test, was also developed. These were administered to a sample of inventory managers at one Army, one Air Force, and one Navy, and two Defense Supply agency installations. The tests are now being scored in preparation for data analysis.

Analysis of the Stability and Instability of Student Growth (887-1)

Mr. Hilton
Mr. Beaton
Mrs. Bower

PURPOSE: To describe the academic growth of a nationwide sample of students over the period from grades 5 to 12--and to investigate the stability and instability in growth of major subcategories of the total sample. This effort is made possible by the Growth Study in which longitudinal data were collected on the achievement, attitudes, opinions, nonschool activities, teachers, schools, and communities of 45,000 U. S. students. The analysis is focused particularly on achievement in mathematics, science, social studies, reading, writing, and listening. Sample subcategories examined included sex, race, early achievement status and growth rate, curriculum, subjects taken, type of school, interests, parents' education and occupation, and school location and size.

PROGRESS: Procedures were developed for fully describing the scores over an eight-year period of 32 subgroups of the total sample categorized by sex, race, father's education, and high school curriculum. Trends over time were analyzed by means of orthogonal polynomials. A final report is in preparation.